ISSN 1018-5593

EUROPEAN COMMISSION

SCIENCE
RESEARCH
DEVELOPMENT

Practical information and Programmes

BRIDGE

Final Report 1994 Vol. II.

EUR 15777 EN

European Commission
Directorate-General XII
Science, Research and Development

BRIDGE

*B*iotechnology *R*esearch for *I*nnovation, *D*evelopment and *G*rowth in *E*urope (1990-1994)

Volume II: Final Report

Edited by
A. Vassarotti, L. Kalbe-Bournonville,
J. Baselga de Elorz and Ph. de Taxis du Poët

1995

EUR 15777 EN

Published by the
EUROPEAN COMMISSION

Directorate-General XIII
Science, Research and Development

L-2920 Luxembourg

LEGAL NOTICE

Cataloguing data can be found at the end of this publication

Luxembourg: Office for Official Publications of the European Communities, 1995

ISBN 92-826-5408-7

Printed in Belgium

TABLE OF CONTENTS

AREA C: CELLULAR BIOLOGY

BRIDGE

N-PROJECTS

AREA A:

INFORMATION INFRASTRUCTURE

PROCESSING AND ANALYSIS OF BIO(TECHNO)LOGICAL DATA

Continuation of the expanded services of the EMBL Data Library (BIOT CT-910254)

COORDINATOR:
G. CAMERON, The EMBL Data Library, Heidelberg, D
PARTICIPANT:
The EMBL Data Library, Heidelberg, D

BACKGROUND INFORMATION

Since the 1970's the determination of nucleotide sequences has become increasingly routine. The EMBL Data Library was established in 1980 ago to build and distribute a database of such sequences, and continues to carry out that work in collaboration with partners in the USA and Japan. By the end of the contract the database was about 120 million base pairs, 200 times the size of its first release in 1982, but less than 5% of the size of a human genome. Advances in sequencing technology and efforts to sequence genomes require constantly developing database methodologies.

OBJECTIVES AND PRIMARY APPROACHES

The goal of the project was to continue to produce the Nucleotide Sequence Database and cope with its rapid growth. These data were to be distributed to users:

(1) by eight quarterly releases on magnetic tape and CD-ROM in the contract period.
(2) using network fileservers allowing access to the latest data over networks.
(3) by daily updates of remote copies at nodes on the EMBnet computer network.

Also during the contract it was planned to improve links between databases and to streamline methods for inclusion of data from large scale sequencing projects.

RESULTS AND DISCUSSION

As anticipated, increased throughput was the most striking development during the contract. Hardware and software developments and refined procedures allowed a threefold speed-up of processing procedures in the first year of the contract. The eight releases planned during the contract were successfully delivered. The database more than doubled in size to about 120 million base pairs during the contract. The growth is summarised in figure 1.

A. Data management

New investments in hardware for the Data Library included the purchase of a database server and a number of workstations. This resulted in savings on software licenses and allowed better user interfaces. This, with software developments using a new version of ORACLE, resulted in substantial productivity improvements. Also much software development was concentrated on submission processing, allowing us to achieve automatic inclusion of submissions created with AuthorIn, and to improve procedures for matching published and submitted sequences.

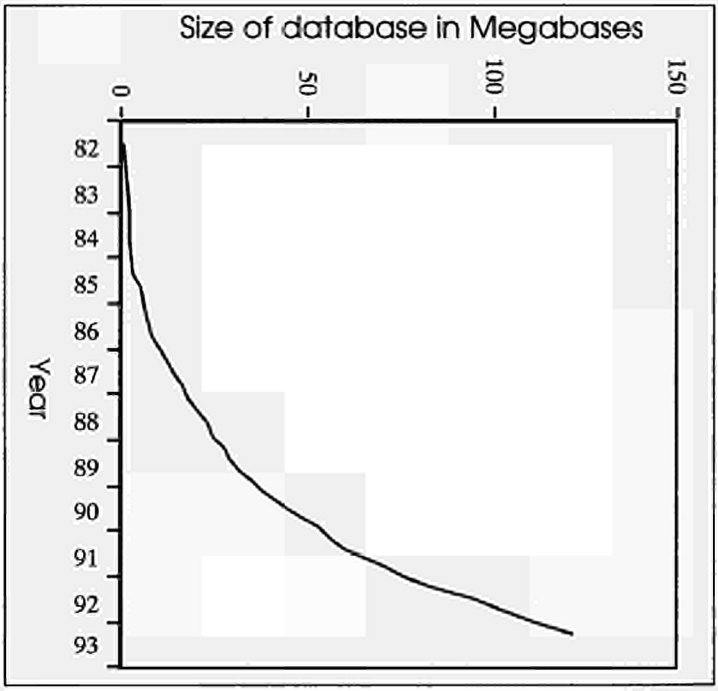


Figure 1. Nucleotide Sequence Database growth

B. Data Distribution

Quarterly releases of the data were distributed on magnetic tape and on CD-ROM, which has overtaken magnetic tape as a distribution medium. New indices allow software to efficiently access the databases on CD-ROM. Also a collaboration with Oxford University Press allowed us to add to our CD-ROM several additional databases described in the Nucleic Acids Research Database Supplement.

C. Network access

The EMBL mail server continued to be popular, with requests for data and programs increasing dramatically. More than 200 programs are now available. A useful by-product of rapid availability of data is feedback from researchers who will, for example, notify us if publications contain data not yet released.

Towards the end of the contract we added a number of new network services to the popular electronic mail server. An anonymous FTP server gives access to complete EMBL databases and software repository including the releases and weekly updates. Also we now offer FTP access via the Gopher protocol which simplifies the use of networks by hiding complexity behind a graphical user interface. Being part of the EMBnet Biogopher network, EMBL's Gopher provides links to other information resources within Europe and elsewhere.

EMBnet continued to flourish with the 15 nodes receiving daily updates. EMBnet is funded under a separate BRIDGE contract.

D. Data from genome projects

During the contract there was a dramatic increase in data coming from genome projects. In collaboration with several European projects the Data Library has implemented procedures to allow these researchers to load data directly into the database. These data are a major factor in the growth of the database. Table 1 summarises the data received from various projects by the end of the contract.

	Entries	*bp*
Genexpress Genethon	3266	998074
C. elegans nematode project	640	408511
S. cerevisiae yeast project	4	327317
French *Arabidopsis* cDNA project GDR	903	271501
Genexpress Munich	239	64870

Table 1. Data from Genome projects

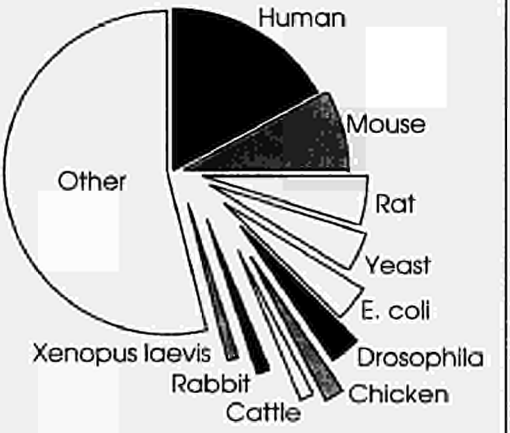


Figure 2. Proportion of Nucleotide Sequence Database from various species.

Although the nucleotide sequence database contains data from more than five thousand species (figure 2), about half the data come from ten species. It is interesting to see how this balance changes over time.

E. Sequence data in patents

Discussions with the European Patent Office (EPO) have resulted in a contract with the EPO providing resources for the database work to be done at EMBL. The backlog of data is made available by EMBL.

F. Research and development

*1. **Identifying near identical sequences*** — The EMBLSCAN tool which was developed in the Data Library has been enhanced for distribution as well as being incorporated into the normal production procedures of the Data Library. This helps to match published sequences and submissions and avoid duplicate entries in the database.

*2. **Multiple sequence alignments*** — The CLUSTAL multiple sequence alignment program was completely rewritten to produce a new program called CLUSTALV. New features include the ability to align old alignments with each other and the calculation of phylogenetic trees.

3. EMBL-Search — A new retrieval system has been developed for the EMBL CD-ROM. It uses index files for lookups of entry names, accession numbers, keywords, species, author names and free text searches. These index files allow any software developer to build such retrieval tools. EMBL-Search is an Macintosh application developed to access the EMBL, SWISS-PROT and PROSITE databases using cross-references to allow users to move between databases to explore related information.

4. MacPattern — The Apple Macintosh program MacPattern for the functional analysis of protein sequences has been improved to support block searches with scoring matrices derived from Henikoff's BLOCKS database and the identification of statistically significant sequence segments according to the maximal segment score method of Altschul and Karlin and, of course, the original Prosite patterns.

5. MacT — We have integrated several algorithms into Apple Macintosh programs for the construction and evaluation of phylogenetic trees.

6. Utilisation of parallel architecture — In a collaboration with John Collins in Edinburgh we provide a service allowing fast database similarity searches using the Smith and Waterman algorithm on EMBL's massively parallel Maspar computer.

MAJOR SCIENTIFIC BREAKTHROUGHS

Improvements in the efficiency of the EMBL Data Library enable it to provide convenient access to a comprehensive collection of data despite challenges from genome sequencing projects.

Other exciting developments have included enhancements to the CD-ROM format and the production of software to exploit them and improved automation of inclusion of data from genome sequencing projects.

The culmination of the deliberations EMBL Council on the future of the EMBL Data Library was the decision to found a new Outstation of EMBL — the European Bioinformatics Institute (EBI) — at Hinxton near Cambridge in the UK. The new Outstation will incorporate and extend the mandate of the Data Library.

MAJOR COOPERATIVE LINKS

International advisory committee — The fourth meeting of the International Advisory Committee for Nucleotide Sequence Databases took place in Washington March 1991, and the fifth in Heidelberg in March 1992. The European members of that committee also met twice during the contract.

GenBank/DDBJ — The collaboration with the GenBank and DNA Database of Japan (DDBJ) groups and data exchange between these groups continued, with two working meetings during the contract. At the first, in Heidelberg, the main topics were the feature table and detailed aspects of annotation conventions. The second, in Mishima, was devoted to data exchange systems being implemented following the transfer of GenBank to the National Centre for Biotechnology Information (NCBI). Also three EMBL staff paid working visits to NCBI to discuss aspects of the collaboration such as editorial standards and data exchange mechanisms.

The Martinsried Institute for Protein Sequences (MIPS) — The collaboration with MIPS, the European partner in the PIR international protein sequence database continued. EMBL forwards protein coding sequences to MIPS and the protein and nucleotide sequence databases pursue a co-ordinated data submission policy.

PUBLICATIONS

Joint publications

Data Library Staff (1991) EMBL Data Library Release Notes and User Manual, Releases **26, 27, 28, 29, 30.**

Bairoch, A. and Data Library Staff (1991) SWISS-PROT Release Notes and User Manual, Releases **17, 18, 19, 20, 21.**

Bairoch, A. and Boeckmann, B. (1991) The SWISS-PROT protein sequence data bank. Nucl. Acids Res. **19**, 2247-2249.

Fuchs, R. and Cameron, G.N. (1991) Molecular biological databases: The challenge of the genome era. Prog. Biophysics Mol. Biol. **56**, 215-245.

Fuchs, R. and Higgins, D. (1991) The EMBL Data Library. In: Collins J. and Driesel A.J. (eds) Genome analysis: from sequence to function. Huethig Buch Verlag, Heidelberg, pp. 149-150.

Pelandakis, M., Higgins, D.G. and Solignac, M. (1991) Molecular phylogeny of the subgenus Sophophora of Drosophila derived from large subunit of ribosomal RNA sequences. Genetica, **84**, 87-94.

Powell, R., Higgins, D.G., Wolff, J., Byrnes, L., Stack, M., Sharp, P.M. and Gannon, F. (1991) The salmon gene encoding apolipoprotein A-I: cDNA sequence, tissue expression and evolution. Gene, **104**, 155-161.

Sharp, P.M., Lloyd, A.T. and Higgins, D.G. (1991) Coelacanth's relationships. Nature, **353**, 218-219.

Waters, A.P., Higgins, D.G. and McCutchan, T.F. (1991) Plasmodium falciparum appears to have arisen as a result of lateral transfer between avian and human hosts. Proc. Natl. Acad. Sci. USA, **88**, 3140-3144.

Bairoch, A. and Data Library Staff (1991) SWISS-PROT Release Notes and User Manual, Releases **22, 23, 24.**

Data Library Staff (1992) EMBL Data Library Release Notes and User Manual, Releases **31, 32, 33, 34.**

Fuchs, R., Rice, P. and Cameron, G.N. (1992) Molecular biological databases — present and future. Trends in Biotechnology **10**, 61-66.

Geber, A., Higgins, D.G., Waters, A.P., Bennett, J.E. and McCutchan, T.F. (1992) Small subunit ribosomal RNA of Blastomyces dermatitidis: DNA sequence and phylogenetic analysis. J. Gen. Mic. **138**: 395-399.

Higgins, D.G. and Stoehr, P. (1992) Fast approximate DNA sequence database searches on compact disc. CABIOS, **8(2)**: 137-139.

Higgins, D.G., Bleasby, A.J. and Fuchs, R. (1992) CLUSTAL V: improved software for multiple sequence alignment. CABIOS **8**, 189-191.

Higgins, D.G., Fuchs, R., Stoehr, P. and Cameron, G.N. (1992) The EMBL Data Library. Nucl. Acids Res. 20 (Supplement), 2071-2074.

Shields, D.C., Higgins, D.G. and Sharp, P.M. (1992) GCWIND: a microcomputer program for identifying open reading frames according to codon positional G+C content. CABIOS, **8(5)**: 521-523.

Individual publications

Bilofsky, H., Informatics in Molecular Biology. 'Medical Informatics Europe 1991' K.-P. Adlassnig, et. al. (Eds.), Springer-Verlag, **28-31**; 1991.

Fuchs, R. (1991) MacPattern: Protein pattern searching on the Apple Macintosh. CABIOS **7**, 105-106.

Fuchs, R., Weber-Schaeuffelen and Zinke, H. (1991) Computer in der Molekularbiologie. In: Bertram, S. and Gassen, H.G. (eds) Gentechnische Methoden. Gustav Fischer Verlag, Stuttgart pp. 307-315.

Stoehr, P. (1991) Sequence databases from EMBL and developments in electronic access. 'Proceedings of the 1991 Chemical Information Conference' H. Collier (Ed), Infonortics Ltd, Calne, 187-195.

Stoehr, P. and Cameron, G. (1991) The EMBL Data Library. Nucl. Acids Res. **19**, 2227-2230.

Higgins, D.G. (1992) Sequence ordinations: a multivariate analysis approach to analysing large sequence data sets. CABIOS, **8(1)**, 15-22.

Luettke, A. and Fuchs, R. (1992) MacT: Apple Macintosh programs for constructing phylogenetic trees. CABIOS **8**, 591-594.

Sibbald, P.R. (1992) Approximate string searching algorithm: technical note. Comm. ACM **35(4)**: 132-137

CarbBank — A structural and bibliographical database for complex carbohydrates (BIOT CT-900184)

COORDINATOR:
K. BOCK, Carlsberg Lab., Dept. Chem., Copenhagen, DK
PARTICIPANTS:
J.F.G. VLIEGENTHART, Univ. Utrecht, Dept. Bio-org. Chem., Utrecht, NL
H. PAULSEN, Univ. Hamburg, Dept. Org. Chem., Hamburg, D

BACKGROUND INFORMATION

The *CarbBank* project is an international effort to supply scientists working in the carbohydrate field with information on complex oligosaccharide structures, mostly of biologically relevant compounds. Besides the three European partners, scientific groups in the USA and Japan are included in the collaboration. The group in the USA, the Complex Carbohydrate Research Center (CCRC) is also responsible for the development of the database management program. The *Complex Carbohydrate Structural Database* (CCSD) is supervised by an international board of overseers, where the coordinator of the project currently is a chairman.

OBJECTIVES AND PRIMARY APPROACHES

The *CarbBank* project is concerned with the establishment of the database (CCSD), containing published oligosaccharide structures higher than disaccharides with biological importance and the accompanying database management program *CarbBank*. Additionally, investigation in connecting spectroscopic information like ^{1}H-NMR and ^{13}C-NMR data to the structural database is made by the establishment of the NMR database *SUGABASE*.

RESULTS AND DISCUSSION

Release number 9 of the CCSD contains 28.800 records of which 11.000 are unique oligosaccharide structures as underivatized trisaccharides and higher. Almost 25.000 records have been examined, about 10.500 by the European partners of the project and the rest by the american partners. Most of the published literature on carbohydrates up to 1993 has been included and the database is almost up to date. The next release CCSD will contain over 30.000 records and be obtainable late 94.

The database and the management program are available either on the Data Repository CD-ROM among other databases distributed by the National Center of Biotechnology Information (NCBI), Washington, DC or on CD-ROM together with the protein sequence database *ATLAS* from the Max-Planck Institute of Biochemistry (MIPS) in Martinsried. Furthermore, the database is accessible over the Internet by anonymous FTP from a public NCBI server.

The main focus of efforts in the period under BRIDGE was to bring the CCSD up to date and to make it as errorfree as possible. Therefore, records obtained from Chemical Abstract Service (CAS) and ISI Current Contents, respectively, were divided under the three European participating groups for verifying against the original literature and correcting, where necessary. Sufficient keywords and additional structures found in the articles were added to the database. A detailed error report was sent to the CCRC for error correction in the central database. Proceeding in this way the three European groups checked over 70% of almost 10.000 records received mainly from CAS during the time under BRIDGE. Over 1.000 additional structures were found in the articles and added to the CCSD

database. Roughly one in three CAS records needed some correction. These modifications were usually small, e.g. the replacement of alpha's or beta's, or misspelling in the bibliographical field. The most frequent transformation that had to be made was the change from free aldose to alditol or pyridylamino derivative. Another amount of CAS records were marked for deletion from the database, most often because the structure set out only was given as carbohydrate composition, e.g. Man5GlcNAc2. This leads to ambiguous structures. In some cases, the structure from CAS could not be found in the original literature. A set of older unverified data (4.029 records) is still in process of verification.

In response to the quite high number of missing structures and errors in the CAS data a special interim report, which documents this in a tabular form and with a few examples, was prepared in Hamburg and will be sent to CAS associated with the current correspondence. The carbohydrate structural and bibliographical data referred to Japanese journals (450 so far) have been extracted and combined into a separated file and will be sent to the Japanese curators for checking together with a detailed instruction list and a few relevant examples.
The NMR database of NMR tables of carbohydrate structures *(SUGABASE)*, developed in the Dutch group, has been extended, and the corresponding management program has been improved. Internally, the program has been divided into separate modules, to make porting to other operating systems and other computer platforms relatively easy. The user interface module has been replaced to enable the portage of *SUGABASE* from MS-DOS computers to the X-Window environment of Silicon Graphics workstations. The database module has been modified to allow the use of the same binary database files on different computer platforms. This allows for one copy of the database files in a network environment. The transformation of the program into a UNIX version for Silicon Graphics workstations is completed and is currently being tested. A manual for this version is almost completed. The manual for the MS-DOS version has been rewritten and examples have been added.
By the end of January 1994 SUGABASE contained 834 ^{1}H-NMR and 380 ^{13}C-NMR records. The program can be obtained as a MS-DOS version and recently as a UNIX version, too. The database and the program are distributed on floppy disks or accessible over the Internet by anonymous FTP from a server at the University of Utrecht.

MAJOR SCIENTIFIC BREAKTHROUGHS

The efforts in building the CCSD have been very successful through the collaboration between the american and european partners; with 28.800 records, the size of the database is larger than anticipated and almost up to date. Actually, the database is well accepted by more than 900 customers worldwide and available not only on CD-ROM but also over the Internet by anonymous FTP. Extra CD-ROMs with the current release will be distributed during the *International Carbohydrate Symposium* in Ottawa in the summer of 1994. With the recently started promising cooperation with MIPS the first step is made to establish the linkage of CCSD to the protein sequence database of MIPS.

MAJOR COOPERATIVE LINKS

The three European partners met every six month to coordinate their efforts, exchange information and evaluate the progress of the project. Meetings with the collaborators in USA and Japan and the established international Board of overseers (K. Bock chairman) took place every second year in connection with the *International Carbohydrate Symposium* or at more informal occasions. Usually, the

communication between the participating European laboratories and with the collaborating groups in USA, respectively is done by electronic mail, fax and phone and works without any problem. Furthermore, K. Bock has been member of the *CODATA* task force on databases in molecular biology to coordinate the activities with other databases.

Nov. 1991 Maastricht (NL) — Symposium 'Bioinformatics in the 90's' and BRIDGE meeting
J.F.G. Vliegenthart, H. Paulsen, A. van Kuik, R. Stuike-Prill, A. Kleen

Feb. 1992 Hamburg (D) — *CarbBank* meeting
K. Bock, J.F.G. Vliegenthart, H. Paulsen, R. Stuike-Prill, A. van Kuik, A. Kleen

July 1992 Copenhagen (D) — *CarbBank* meeting
K. Bock, R. Stuike-Prill, A. van Kuik, A. Kleen, S. Doubet (CCSC), B. Bossenbroek CAS)

July 1992 Paris (F) — Board of Overseers
next to the XVIth International Carbohydrate Symposium
K. Bock, J.F.G. Vliegenthart, H. Paulsen, P. Albersheim (CCRC), S. Doubet, T. Ogawa, B. Bossenbroek

Nov. 1992 Hamburg (D) — small *CarbBank* meeting
K. Bock, H. Paulsen, A. Kleen

Nov. 1992 Utrecht — small *CarbBank* meeting
K. Bock, J.F.G. Vliegenthart, A. van Kuik

May 1993 Copenhagen (DK) — *CarbBank* meeting
K. Bock, J.F.G. Vliegenthart, H. Paulsen, R. Stuike-Prill, A. van Kuik, A. Kleen

Nov. 1993 Utrecht (NL) — *CarbBank* meeting
K. Bock, J.F.G. Vliegenthart, H. Paulsen, A. van Kuik, A. Kleen

PUBLICATIONS

Joint publications

CarbBank and the Complex Carbohydrate Structure Database. R. Stuike-Prill, K. Bock, A. Kleen, H. Paulsen, J.A. van Kuik, J.F.G. Vliegenthart, S. Doubet, D. Smith and P. Albersheim, *Bioinformatics* (1992) 1: 12-15

Studies on the carbohydrate chains of glycoproteins. J.F.G. Vliegenthart and Conformational Assessment of Glycopeptides, K. Bock, R. Stuike-Prill and M. Meldal. Alfred Benzon Symposium 36. In: K. Bock and H. Clausen, *Complex Carbohydrates in Drug Research, Structural and Functional Aspects, 1994.*

Individual publications

CarbBank version 2.2 with the database CCSD.5 (1991)

A ^{1}H-NMR database for complex carbohydrate structures. J.A. van Kuik and J.F.G.Vliegenthart, *TIGG* (1991) 3: 229-230

CarbBank version 2.3 with the database CCSD.6 (1992)

Databases of complex carbohydrates. J.A. van Kuik and J.F.G. Vliegenthart, *Trends Biotechnol.* (1992) 10: 182-184

A ^{1}H-NMR database computer program for the analysis of the primary structure of complexcarbohydrates. J.A. van Kuik and J.F.G. Vliegenthart, *Carbohydr.* Res. (1992) 235: 53-68

CarbBank version 2.4 and the databases CCSD.7, version 2.5 with CCSD.8 (1992/1993) and version 2.6 with CCSD.9 (1994).

Databases of complex carbohydrates. J.A. van Kuik and J.F.G. Vliegenthart, *Trends Food Sci. Technol.* (1993) 4: 73-77

A NMR spectroscopic database of complex carbohydrate structures. J.A. van Kuik and J.F.G.Vliegenthart, *Biotech Knowledge Sources* (1993) 6: 3
The structure of the carbohydrate backbone of the core-lipid-A region of thelipopolysacccharide from *Vibrio Cholerae* strain H11 (non-O1). E.V. Vinogradov, R. Stuike-Prill, K. Bock, O. Holst and H. Brade, *Eur. J. Biochem.* (1993) 218: 543-554

Protein Sequence Databank (BIOT CT-900170)

COORDINATOR:
H.W. MEWES, Max-Planck-Institut f. Biochemie, Martinsried, D

BACKGROUND INFORMATION

Protein sequence data are of major importance for basic research in the biomedical sciences as well as for biotechnological infrastructure. The collection and distribution of protein sequences in Europe has been the subject of our work. MIPS is the European contributor to the international network of protein sequence databanks. The growing data streams from large scale genomic sequencing projects will further increase the work required for comprehensive annotation of the sequence data. Structuring the database into protein families and developing new concepts for organizing biological data will improve the efficiency of the annotation process significantly and will increase the usefulness of the database to the general scientific community.

OBJECTIVES AND PRIMARY APPROACHES

The unique goals of the Protein Sequence Database are:

1. to dynamically maintain a complete and comprehensive set of published protein sequences and related information in accordance with current biological understanding
2. to organize these data by similarity and evolutionary relationship
3. to add biochemical information that is displayed in a consistent and comprehensive form.

The Protein Sequence Database is now capable of effectively managing the increasing data flow from the literature and the submissions to the nucleic acid sequence databanks. Currently more than 70,000 protein sequences are represented in the database, approximately 50% of the worldwide data input is provided by MIPS.

RESULTS AND DISCUSSION

The Martinsried Institute for Protein Sequences (MIPS) at the Max Planck Institute for Biochemistry has developed resources to support data collection and distribution as part of the tripartite collaboration of the Protein Sequence Databanks (PIR-International). The work, funded by the BRIDGE project, concentrated on the development of the Protein Sequence Database, particularly on the standardization of its annotation and on its distribution on networks.

Sequences in the Protein Sequence Database are derived from publications in books or scientific journals or from author submissions to the sequence databases.

(i) Database Progress

During the report period, the PIR-International Protein Sequence Database increased by 242% to 64,760 sequences by release 39.0, Dec. 1993. The growth is exponential with a doubling time of 2.4 years. This contrasts to doubling times of 3.5 years for both SwissProt and SEQDB (Protein Research Foundation, Osaka) during the same period. PIR-International is both the largest and fastest growing

protein sequence collection available. These figures testify to the success of the project. A major effort focused on the development of annotation strategies, the key problem of sequence data processing. A formally-defined conceptual database schema (as represented by a revised format, CO_2), designed to facilitate consistent full annotation[2], was developed and the initial steps for implementation of an object-oriented database system have been taken. The new schema and the incorporation of database components into an object-oriented database system will allow the implementation of a truly distributed database operating between the centers in Washington and Martinsried; eventually, it will be extended to the node in Tokyo.

Services provided have been extended. The database is currently distributed on CD-ROM for VMS, Alpha-VMS, ULTRIX, SUN-OS, Macintosh and MS-DOS systems, including the ATLAS software for concurrent access to multiple databases.

(ii) A database of sequence similarities

The development of an exhaustive database of sequence similarities (FASTA database) was completed. Concurrent updating of this data set and the Protein Sequence Database allows immediate access to the closest relatives of any sequence in the database. The FASTA database has proven to be extremely useful: both for the scientific evaluation of the data and for daily use by the annotation staff. It serves as a basis for the classification of sequences and for their annotation. It also provides a facility for unbiased statistical analysis of database growth.

Every protein sequence added to the sequence collection is compared with all other sequences available and the results are stored. While it is possible to run such comparisons on demand, without a mechanism to ensure concurrency, the results become outdated quickly and of limited value. The FASTA database system solves this problem by saving the results in compressed form and incrementally updating them as new or modified data are incorporated into the sequence collection. The database is updated weekly and currently contains (March 1994) 140,000 entries. FASTA results are also stored for preliminary data in preparation for inclusion into the PIR-International database.

(iii) Progress in the annotation of protein sequences

Standarization of data and tools to control data processing and data flow are required to maintain correct, complete, and comprehensive data sets, such as the protein sequence database.

A comprehensive protein sequence database must provide
- reliable sequence data
- complete coverage of available protein sequence data
- high quality annotation: the biological information associated with a protein sequence should be complete and consistent
- organization of the data: data should be represented in a well-defined form that is suitable for database operations

Defined lists of terms were compiled and their usage in the database was made consistent with these lists for the following classes of data:

Species names
The taxonomy list, is applied to the source field of the database entries.

Enzyme names
The nomenclature of the Enzyme Commission is employed. The correspondence between EC number and 'recommended protein name' is monitored for all entries in the database.

Non-enzyme protein names
In contrast to the enzyme names, no nomenclature for non-enzyme proteins exists. To achieve standardization for non-enzyme protein names, lists of approved, undecided, and inadmissible names have been compiled.

Keywords
A list of selected keywords has been developed. Major releases of the database contain only valid keywords. Keywords yet missing in entries are continuously being added. Since keyword lists do not describe the application of individual keywords, a set of rules is being developed to describe the relation between the set of defined keywords and the properties of proteins.

Superfamily names
The list of superfamily names has been generated and all entries are checked for the occurrence of new or invalid superfamily names. Major releases contain valid superfamily names.

Features
All features have been standardized during the report period. To facilitate the annotation of features, the database is being organized into protein families using multiple sequence alignments. Directed efforts to spread features, keywords, and other associated information among homologous proteins will be initiated shortly.

(iv) Progress in protein classification

One limitation in using the present PIR database is the lack of a complete sequence classification: only 12% of the sequences contained in the Protein Sequence Databank have been classified formally. The remaining sequences are gathered in PIR2 (partially annotated) and PIR3 (preliminary entries). The artifical separation into the datasets annotated and unannotated entries will be removed.

To achieve this goal two sets of procedures have been developed:

— Semi-automatic introduction of multiple sequence alignment and sequence database search results into the PIR-International protein sequence classification

— Automatic classification of sequences according to established classification rules56% of the dataset have been classified.

(v) Data distribution on wide area networks

In 1992, MIPS began contributing actively to the EMBnet. Established as a special node, MIPS forwards incremental updates of the protein sequence database to the network of European national nodes.

There are two important considerations in the development of an effective strategy for large-scale database access.

— The scale of the demand for access to macromolecular sequence databases places severe loads on computer and network resources at centralized sites.

— The user community exhibits a broad range of needs that will not likely be satisfied by a single centralized site that can provide only a limited class of database access interfaces.

MIPS has developed a prototype system for propagating incremental database updates, with complete synchronization, across a network of *satellite sites* to address this problem 4-6. The remote satellites are organized in a directed-graph topology that can be optimized to balance the load on any particular node or network link in the configuration. Multiple propagation paths buffer the system against individual nodal failure and preserve the integrity of the global network. The distribution protocol is independent of the database application running at individual nodes; hence, the distribution system can support a diverse set of remote sites providing a wide variety of different services. Because complete synchronization is assured, nodes can operate cooperative, compatible services. For example, one site may operate a specialized sequence database or pattern recognition service while another may provide a textual query language interface. Users can access both sites and in combination benefit from both broad ranging and specialized services.

This capability is made possible by adopting a layered architecture that insulates the network, data distribution, and application software layers from one another making them formally independent. Application layer software developers may focus exclusively on the application, while being guaranteed complete and reliable access to the data store.

MAJOR COOPERATIVE BREAKTHROUGHS

Regular production of a timely comprehensive data set of protein sequences on CE-ROM, magnetic media and electronic networks.

MAJOR COOPERATIVE LINKS

A number of cooperative efforts have been undertaken with European labs, partially in conjunction with the BRIDGE Yeast Sequencing Project. MIPS serves European laboratories by providing the most up-to-date protein sequence data collection, distributing that data, providing training in sequence data analysis, and providing user support. As part of a collaboration with the EMBL Data Library, the latest nucleic acid sequence data collected are forwarded to MIPS.

PUBLICATIONS

Barker W.C., George D.G., Mewes H.W. and Tsugita A.; 'The PIR-International Protein Sequence Database', Nucleic Acids Res. (1992) **20**: 2023-2026

Barker W.C., George D.G., Mewes H.W., Pfeiffer F., Tsugita A.; 'The PIR-International database', Nucleic Acids Res. (1993) **21**: 3089-3092

George D.G., Orcutt B.C., Mewes H.W., and Tsugita A.; 'An Object-Oriented Sequence Database Definition Language (SDDL)', Protein Sequences and Data Analysis (1994) **5**: 357-399

Heumann, K.; 'Spezifikation und Implementierung einer Datenverteilinfrasturktur für Proteinsequenzdatenabanken', Diploma Thesis, Univ. of Munich (1992)

Heumann, K., Mewes, H.W., and George, D.G.; 'A new concept of sequence data distribution on wide area networks', Computer Applications in the Biosciences, submitted, (1994)

Mewes, H.W., Doeltz, R., and George, D.G.; 'Sequence databases: an inevitable source for biotechnological research', Journal of Biotechnology (Special issue on genome research/ molecular biotechnology), in press, (1994)

The promotion of EMBnet: computer network for bioinformatics in Europe (BIOT CT-910273)

COORDINATOR:
C. SACCONE, CSMME-CNR, Bari, I
PARTICIPANTS:
H. ULLITZ-MÖLLER, Aarhus University, DK
J.C. ROOS, University of Helsinki, SF
G. CREMER, Université René Descartes, F
S. SUHAI, DKFZ, D
A. BLEASBY, Daresbury Laboratory, GB
C. SAVAKIS, Inst. of Mol. Biol. and Biotech., GR
J.H. NOORDIK, University of Nijmegen, NL
E. PAULSSEN, Biotechnology Centre Oslo, N
J.M. CARAZO, C.S.I.C., E
G. MAGNUSSON, Uppsala Biomedical Centre, S
R. DÖLZ, Biozentrum Basel, CH

BACKGROUND INFORMATION

Bioinformatics is a rapidly growing science and an ever wider variety of skills and expertises (scientific and technical) are needed to assure Europe is competitive and well served in this field. EMBnet (European Molecular Biology Network) was constituted in 1988 during a meeting at the EMBL Data Library in Heidelberg with the purpose of creating an infrastructure for molecular biology-based bioinformatics in Europe, including both service and research activities. It was established as a computer network for, and formed by, the European molecular biology and related research communities. It would provide an international support and interaction group for national and specialist nodes which themselves serve different communities of European researchers. EMBnet, presently draws on 24 centres, which accounts for a wide net of talent upon which to rely.

EMBnet was indeed conceived to be a 'laboratory without walls' where the various skills available in Europe would come to profit the entire bioinformatics community. As such it accommodates the different demands and needs of the national and scientific communities it represents. By its physical presence across the community, it provides for rapid interaction with users, in their own language and against their own local background, but it also concentrates these opinions into common policies. In addition, EMBnet is an evolving service, with its scientific focus centred on, but not solely confined to, providing databases, software and related technical and educative support to nucleotide and protein sequence and structure databases. EMBnet is thus able to serve the rapidly expanding demands on bioinformatics, and can, through its diversified structure, react to accommodate these changes as they occur by attracting new nodes or linking with other services as the needs arise.

As EMBnet's users require more than EMBL Data Library output alone, the group is already exploring how best to offer other data services to its members. EMBnet users require closer ties to expertise present at each node which can help them make the best use of the international array of data now available.

OBJECTIVES AND PRIMARY APPROACHES

Against this background, obtaining the data is perhaps no longer a major problem, but the data has to be linked, manipulated and integrated into R&D targeted projects and the closer the ties between experts and user the better. EMBnet seeks to provide the total European Bioinformatics Community with a technical and intellectual infrastructure, access to biological data and programmes, and support for biocomputing research. EMBnet endeavours to ensure that national nodes, receiving data from commonly agreed international sources, operate in harmony to make this information available to their national users in the best available manner. The experiences and benefits gained in one country can thus be enjoyed by others.

EMBnet works through training, executing agreed R&D, and informing new users through promotional tools of what the field needs and has available. All these activities are carried out by teams drawn from EMBnet nodes devising, bidding for, judging on and executing, defined programmes and projects. To formalise the procedures used in judging such projects, the group produced their own 'Guidelines for Funding Proposals' which has led to a secure peer-reviewed infrastructure structure.

Initially the total programme centred upon improving the transfer and access of the databases produced by EMBL. Then, as more general data and databases appeared, and new problems facing the use of the data emerged, new project areas were developed. In view of this 'shifting scenario' EMBnet did not lay down hard and fast rules for research areas. 35 titled projects (see Tab. 1) were carried out and their major achievements are reported in *'Major scientific breakthroughs'*. The results of these activities have formed the basis for better general EMBnet services. Not all of these are 'node-dependent' and many have a far reaching impact on the bioinformatics communities outside EMBnet. EMBnet's 'Institute without Walls' structure allows indeed questioners from anywhere within the molecular biology community to call any node and ask for support and advice.

EMBnet's proven goal is to provide an infrastructure for, especially, molecular-biology-based bioinformatics in Europe. To continue to do this, EMBnet requires a formal structure. A number of long-term legal structures were researched before the membership agreed, at the 7th Meeting Basel, October 1993, to become a Stichting registered under Dutch law. This registration took place on March 17th 1994.

RESULTS AND DISCUSSION

For the purposes of this report, EMBnet's activities have been divided into: R&D, Training, and Organisation.

A. The R&D Programme:

Bioinformatics is a young and evolving discipline; and a large number of 'problems' and 'needs' appear in relation to daily experience. Initially, the programme centred upon improving the transfer and access of the databases produced by the EMBL Data Library. As the programme progressed, other projects were added to the activities. Generally, as more general data and databases appear, and more detailed coordinates and measurements of the subject matter is achieved, clearer paths of applied research with regard to the better utilisation of these data become visible.

EMBnet project appraisal allows submissions in a variety of areas, as shown in the list of projects reported in Table 1. The descriptions of these projects have to be read with the time frame in mind — many projects were started early in the grant period and led to further activities. The EMBnet membership is very pleased with the results attained: all the accepted projects were completed successfully.

The results of these activities have formed the basis for better general EMBnet services. Not all of these are 'node-dependent' and many, such as Network Data Transfer Protocol (NDT), the Sequence Retrieval System (SRS), and the Extended EMBL Keyword project have a far reaching impact on the bioinformatics communities outside EMBnet.

EMBnet is clearly a collection of nodes each of which has a national or special mandate. As the programme matured, it became clear that the group could share resources as well as expertise. The EMBnet User Interface Menu System and the establishment of a central service (the Biocellerator) for all EMBnet nodes to access and test were first steps in what is expected to be a continuing plan to make even better use of national resources at a European level.

B. Training:

Any new discipline requires training. Bioinformatics, with its reliance upon computer and information technologies is especially dependant upon well trained personnel — often across disciplines — and many nodes give a series of courses based upon their own services.

The EMBnet nodes hold local courses according to their own schedules. Some countries, such as Denmark, hold these on a bi-weekly schedule and even the small, young nodes, can show an impressive number of students attending (e.g. the Belgium node, in its initial start-up phase, has been training more than 100 students per annum and this number is set to increase).

The EMBnet's training committee has built upon this core activity and organised a large number of international bioinformatics courses. These have usually been held in a national node, often based at least partially in the local language. A basic EMBnet Bioinformatics course, aimed at 'starting students in bioinformatics' has also been developed whose structure can be used by other nodes anxious to start users off on the right foot. This course, consisting of 2 weeks intensive hands-on training, has been held twice in conjunction with ICGEB at their node in Trieste Italy. In addition to these core courses, specialist courses are increasingly essential. These have varied from 'The use of Internet' to 'The use of specialised software in molecular modelling'. In the last few months of the programme the EMBnet Spanish node ran a course with the support of EMBnet members and money which was advertised internationally, as the awareness of the field and the role of EMBnet within it (for the full list of Training courses, see Tab. 1)

To maintain quality and to identify areas for improvements, all courses have been critically evaluated by both the givers and the takers, and are being constantly improved. These activities have shown that a number of EMBnet personnel are excellent educators.

C. Organisation:

(see Tab. 1 for relevant funded actions).

Table 1: EMBnet activities funded through BRIDGE grant

Date	*Title of proposal*	*Proposer*	*Funding area*	
1.	March 1992	Hiring of a manager	Scientific coordinator	Management
2.	January 1993	EMBnet/BRIDGE brochure	Manager/ Coordinator	Management
3.	January 1993	Project manager	Steering Committee	Management
4.	December 1993	Public Relations and Promotion of EMBnet	Steering Committee	Management
5.	December 1993	Establishing EMBnet as a Stichting	Steering Committee & Coordinator	Management
6.	October 1991	Three workshop meetings on conferencing and access to remote services	EMBL	Meeting
7.	March 1992	Annual EMBnet business meeting	Dutch node	Meeting
8.	June 1992	Travel grant for Network Services Conference	EMBL	Meeting
9.	June 1992	Travel expenses of the Steering Committee	Manager	Meeting
10.	October 1992	EMBnet meeting	Portugese node	Meeting
11.	September 1993	EMBnet meeting	Swiss node & Hoffman-La Roche — special purpose node	Meeting
12.	December 1993	Node visits	Greek node	Meeting
13.	September 1993	Travel grant for Internet Engineering Task force	Swiss node	Meeting
14.	October 1991	Networking data distribution	EMBL — special purpose node	Networking
15.	October 1991	Acquisition of Multinet TCP/IP software	Swedish node	Networking
16.	March 1992	Investigation and prototype development of 'next generation' sequence database	Swedish node	Research & Development
17.	March 1992	Interactive remote access to high performance computing resources for sequence analysis	Edimburgh, UK node, EMBL — special purpose node	Networking
18.	June 1992	Protein sequence data distribution on Wide Area networks	MIPS — special purpose node	Networking
19.	June 1992	Contribution for network services	Swiss node	Networking
20.	June 1992	EMBnet menu system	Dutch node	Technical development
21.	October 1992	Gopher access to FTP sites	Finnish node	Networking

22.	October 1992	Release of data manipulation software	Swiss node	Research & Development
23.	January 1993	Extension of the scope and query capability of SRS	EMBL — special purpose node	Research & Development
24.	January 1993	Setting up TCP-protocol transfer to Heraklion	Swedish & Greek nodes	Technical exchange
25.	December 1993	WWW/XMOSAIC Project	Finnish node	Research & Development
26.	December 1993	NDT protocol on nodes not yet included in network	Swedish nodeResearch & Development	
27.	December 1993	EMBnet Bioaccelerator Services	Dutch node	Research & Development
28.	December 1993	HASSLE protocol on nodes not yet included in network	Swiss node	Research & Development
29.	December 1993	Running HASSLE on PCs	Norwegian node	Research & Development
30.	March 1992	Genes, proteins and Computing conference	UK node	Training
31.	June 1992	Practical course on genome information	EMBL — special purpose node	Training
32.	October 1992	Visiting scientists	Dutch node	Training
33.	October 1992	Proposal to make a place available to an EMBnet user with travel and accommodation funded on the computing courses of the UK Human Genome Mapping Project.	HGMP — special purpose node	Training
34.	October 1992	Madrid training courseSpanish node	Training	
35.	October 1992	Proposal for 5 day course on molecular sequence analysis to be held at Istituto di Tecnologie Biomediche Avanzate, CNR, via Ampere 56, 20131 Milano, Italy, 14-18 December 1992	HGMP — special purpose	Training
36.	October 1992	Vienna training course	Austrian node	Training
37.	January 1993	Trieste training course	ICGEB — special purpose node	Training
38.	December 1993	Basic User Training Course Trieste and Knowledge Exchange	ICGEB — special purpose node	Training
39.	December 1993	Madrid Training course 'Frontiers in Bioinformatics'	Spanish node	Training

EMBnet has undergone a series of changes during its 5 years of existence. These changes have taken place as the group has matured, and seen fit to define its own

role and structure. EMBnet, at the start of this BRIDGE project, was a collection of nodes doing little more than receiving and making available EMBL data. The BRIDGE grant gave the group the opportunity to develop further, to examine what the needs of the community were and how these could be met.

At the start of this grant, EMBnet organised itself under a 'Consortium Agreement' agreeing to support the BRIDGE project and each other through a Letter of Agreement issued by the Scientific Coordinator of the project. In this structure EMBnet was still, in fact, only a 'transient grouping' but, during the first two years of the grant, it became evident that a clear role for EMBnet had evolved and that a more formal, long-term, structure allowing EMBnet to become self-governing and fund-seeking was required.

It was also clear that EMBnet would, and should, continue to grow. New nodes, new users, and new expertises were and are joining the user community. The BRIDGE project had clearly illustrated the need for more formal management and administration and so, in 1992, the group elected a Steering Committee to assist and support the Scientific Coordinator in the daily running of the project. This group consisted of Alan Bleasby (Daresbury Laboratory), Robert Harper (University of Helsinki), Jan Noordik (CAOS-CAMM, University of Nijmegen) and Chris Sander (EMBL), and was assisted by a part-time executive manager (J. Franklin (Manager of ASFRA BV taking over from A. Kroon (AL BV, Clinical test Manager, who had assisted C. Saccone until December 1992).

Starting at the 5th EMBnet Business Meeting, a number of long-term options for forming a stable, legal, structure were initiated. Options such as a Company limited by Guarantee, a European Economic Interest Group, and a Foundation, in the form of a Dutch Stichting were researched and the membership agreed, at the 7th Meeting Basel, October 1993, to adopt the latter structure. The last six months of the grant period have resulted in the legal papers being drawn up and agreed so that, as of March 1994, EMBnet is legally known as The EMBnet Stichting.

In parallel with these developments, the membership had organised itself into a series of Working Committees, with the Steering Committee, aided by an Executive Manager, being responsible for the daily execution of EMBnet's activities. The direction and form of EMBnet is decided by the full membership, currently 24 nodes, meeting annually to decide such issues. Each committee is responsible for a set number of tasks or actions.
Some are ad hoc, for instance to prepare the content of the EMBnet Brochure, but others, such as the group dealing with Training, are semi-permanent and consist of members who have professed an interest in the subject concerned.

EMBnet maintains a number of electronic news and information groups. The node managers are formally linked through the EMBnet ADM grouping and, with the formation of the European Bioinformatics Institute, a new grouping EMBnet-EBI has been formed. A number of EMBnet members monitor and select, and remount, material from the international groups. In this way the group is kept fully up-to-date with the field and with relevant developments.

D. Conclusions

EMBnet will continue to do what it does well. Plans for further training, for applied R&D, for better promotion and for the further support of bioinformatics in academia and industry are well advanced. Discussions with the EBI (whose recent statements make it clear that they, too, see a complementary role for the two organisations) are under way to establish even better collaboration and service in the coming years. Discussions with other database producers are taking place.

EMBnet nodes are also supplying staff to support and guide the many committees and related activities covering European networks. The biology market is now a major network. EMBnet has established contacts with he various European and International organisations examining these needs.

Now that EMBnet has a formal structure, it will actively seek new members and new outlets for information.

The BRIDGE project 'The promotion of EMBnet: computer network for bioinformatics in Europe' has, in our opinion, succeeded. EMBnet has a structure, has produced many excellent solutions to perceived needs and is now ready to coaccessible 'pool of talent' for the future.

MAJOR SCIENTIFIC BREAKTHROUGHS

Seven funding calls were held throughout the grant period to select R&D proposals at the various nodes. All proposals were subjected to a refereeing procedure before a decision was made. Many projects were accepted after revision.

The following projects can account for major achievements in the course of BRIDGE grant in the light of their contribution to the development of bioinformatics.

A. Design and development of an improved mechanism for data exchange on EMBnet based on RFC977 (NNTP).

The proposal financed a contract between the EMBL node on EMBnet and a computer scientist/programmer, to improve the existing mechanisms for data transfer on EMBnet; mainly concentrating on the exchange of sequence data.

The transmission of new sequence data from EMBL to the national nodes each night via DECnet/X25 and TCP-IP/Internet is at present the main endeavour of the EMBnet project, and therefore reliable data transfer is crucial for the success of EMBnet in general and for the national nodes if they are viewed as remote copies of the EMBL database. The existing procedures which control data transfer were developed at EMBL. Although they include simple checks for successful data transmission, they cannot fully guarantee to synchronise the data collections at EMBnet nodes with the master data collection at EMBL, and a fair amount of human intervention is required to keep the copies of the database in sync. The current procedures do not propagate deletions of entries and entry name changes, and they cannot efficiently cope with poor connections. Additional complications arise from the fact that the EMBnet concept envisages the redistribution of data to customers of the national services.
The current procedures are too specific to support this notion. Database entries can be seen as a special kind of 'news items', which is at least true for the current flat file format. A standard already exists for the exchange of news items: the Network News Transfer Protocol (NNTP), described in RFC977, is the de facto standard for transmitting news items on the Usenet logical network. The NNTP protocol describes how server and clients cooperate to keep lists of news items in sync at remote places.

Thus, NNTP seems to be a good starting point for the development of an EMBnet data transmission protocol.

Genbank has recently started to exploit the NNTP protocol and the Usenet news system for the distribution of new sequence database entries. In contrast to the EMBnet philosophy, it is the responsibility of the recipients of GenBank sequence news items to use the received messages to keep their local database in sync with

the master database. Therefore, for EMBnet purposes it does not seem to be appropriate to simply copy this approach and to create kind of 'EMBL newsgroup'. Instead, research is necessary to solve the problems of synchronising the master database at EMBL and the remote copies at the EMBnet nodes.

The NNTP protocol is currently mainly used in conjunction with newsreader software such as ANU NEWS, news, etc. The functionality of these programs is not required for an EMBnet data exchange mechanism. The basic requirements for a server and a client (which of course do not have to be separated programs) are to keep two lists of database entries at remote sites in sync and to deliver newly received data in an appropriate form. In view of the diversity of hardware platforms used by the EMBnet nodes it is desirable to write code which is not dependent on operating system-specific features. Sequence data received by national nodes are further manipulated in very different ways, depending on the local database management and sequence analysis software which makes use of it. It is, therefore, necessary to develop standards for interfacing the EMBnet data distribution mechanism to software which acts on the delivered data.

B. Interactive Remote Access to High Performance Computing Resources for Sequence Analysis

The Biocomputing Research Unit at Edinburgh University has developed an extensive series of programs for sequence database analysis, making use of the AMT Distributed Array Processor ('DAP'; see, e.g., Collins,J.F. and Coulson, A.F.W. 'Significance of Protein Sequence Similarities' in Methods in Enzymology (1990) 183, pp 474-487). The DAP provides the most cost-effective architecture for many sequence analysis algorithms, and we believe that the DAP is the only machine world-wide on which exhaustive searches of the entire sequence databases using the Needleman-Wunsch-Sellers algorithms are run routinely. DAP's running this software have also been installed at ICRF in London and as part of the computing resource of the UK Human Genome Project. Versions of these programs can be run remotely by dispatching electronic mail in a simple format to an automatic mail server process on the Edinburgh Sun which hosts the DAP, and this facility is used by a large number of groups in the UK (under the auspices of the UK EMBNet node at Daresbury), elsewhere in Europe, and in the rest of the world. However, this mode of access is non-interactive and can practicably be made to support only the simplest applications of the database search software. We are now seeking support for the development of an interactive front-end to the DAP facilities in order to allow access to the full power of these programs from the EMBNet area Ideally this front-end should provide a client for remote X-windows servers. We anticipate that this should now be straightforward to implement within the UK but our commitment is to provide a system which will be accessible very widely within EMBNet. A preliminary study will therefore be needed to establish whether this level of interaction is practicable throughout Europe, and if it is not, to specify software which can be used.

The preliminary study and software specification will be made by Andrew Coulson and John Collins, in consultation with Alan Bleasby and Peter Stoehr. If necessary, visits will also be made to other EMBNet sites in this phase, which will be completed by June 1992. Software specified will be written under contract by a graduate computer scientist/programmer, in a period of up to eight weeks in the summer of 1992. This work will be done under the supervision of Andrew Coulson in Edinburgh. Daresbury(Alan Bleasby) and EMBL (Peter Stoehr) will act as test sites for the software. Training in the use of the software will be provided by visits

by Andrew Coulson or John Collins to appropriate EMBNet node sites in the second half of 1992 and early in 1993.

C. Investigation of design and prototype development of the 'next generation' sequence database.

The current generation of sequence databases are based on files, not databases. There are several reasons for this: first, application programmers have not had access to expensive database software, and secondly, biological data does not fit very well into record-oriented models of relational database management systems. Relational databases are also too slow for interactive design applications that need to extract a hierarchy involving many joins. There has also been little motivation for concurrency and shared data in the past. The next generation databases will need to support both team productivity, collaborating authoring, the speed of interactive applications and also some sort of standardization.

With the emergence of object database technology, it has become possible to handle the data types biological database applications require, and to do this with a performance comparable to that of their single-user file based predecessors.

It is important that access to these databases can easily be performed with languages such as Fortran or C and not only with object oriented languages such as C++ or Smalltalk.

For many years, EMBL has had a relational database which they have retained for internal use only. Genbank, meanwhile, has developed a relational database which they are currently distributing to satellite sites all over the world with daily updates. Uppsala is one of the first Genbank satellites in Europe. This now complements the data distribution which for many years has occurred by EMBnet.

Dr. Rainer Fuchs has received funds from EMBnet to investigate the next generation of data distribution protocols. It was agreed that a general protocol to transfer object oriented data should be investigated.

There is no need for EMBnet to invest resources in developing a new relational database system, but instead focus on what comes next. We think that this fits very well in with Dr Fuchs project and eventual extensions from it. After the investigatory phase, the implementation can easily be arranged as a collaborative project.

There is of course the possibility that a hybrid relational/object oriented approach would be the most powerful, although the cost of software may be prohibitive.

D. Collaborative Project for the Design and Development of an EMBnet User Interface or Menu System

EMBnet nodes have been defined to fulfil requirements on database access, network services and data distribution and user support. This last task has been very well documented by Dr. A. Bleasby et. al in a SERC document 'Improving Front-line User-Support: Proposals for Greater Efficiency, in Computational Molecular Biology', 1991.

In this document three major problem areas facing molecular biologists using computational methods are identified.

1. The difficulty of program choice owing to the fast- changing, somewhat structureless array of software from which to select.

2. Confusion in executing the software: this is partly due to the low priority given by some program writers in providing intuitive program interfaces.
3. Difficulties encountered in specifying program parameters.

These often reveal an insufficient understanding by the user of how to apply a method, i.e., the molecular biology, embodied in the program. Particularly the first two problems, and to some extent the third one, can be solved by providing the molecular biologist with intuitive interfaces to the software and the databases and to the (software) utilities needed to provide network services. We propose to develop such an interface; a General Menu System.

A recent survey of end-user equipment of molecular biologists in The Netherlands has shown (and one expects a similar situation at end-user sites of other Regional EMBnet nodes) that a vast majority of molecular biologists is still using character oriented hardware as a terminal facility (mainly of the PC/emulator type with serial line connections). Despite the proliferation of workstations (including ether-netted PC's or Macintoshes) this situation is not expected to change drastically in the coming 4 year period. Moreover, wide area network performance (over a whole country or region) both at the national and the European level in EMBnet, is currently not such that it can support 'massive' graphics oriented (e.g. X) data interchange between end-users and Regional EMBnet nodes, ruling out the general usage of graphical user interfaces.

The problems faced by the molecular biologist with his terminal hardware and confusing array of programs and databases, as well as the network performance, urge the development of a (character based) general user interface. The importance of such an interface has since long been recognized among the EMBnet nodes. Appendix B-3 of the Crete Meeting (May 10/11, 1991) prepared by Dr. R. Doelz summarizes in paragraph 3.2.3 the situation at that point in time and action is suggested but has never been taken in a collaborative fashion. Now (Dec. 1992) the conclusion of paragraph 3.2.3. of the report can be re-phrased into: *Menu systems based on SMG$ (VAX VMS) and CURSES (Unix) CAN efficiently be used on most IP connections in EMBnet. Graphical User Interfaces (GUI's) based on the X-Windows Interface CANNOT currently be used on these connections.*

To improve user support at EMBnet nodes, this proposal describes a collaborative project for the development of an EMBnet GENERAL MENU SYSTEM (EMBnet GMS) which will cover the major EMBnet node software functionality. In particular emphasis should and will be placed on guided access to multiple program collections and databases, to networking navigation tools like Gopher, WAIS e-mail, NEWS etc., and to intelligent HELP for molecular biology application programs.

This has been a major project seen as key to the future unfolding of EMBnet as a series of nodes that can interact and support each other, possibly evolving to the point where users transfer 'transparently' across the network.

E. Extension of the Scope and Query Capability of SRS.

SRS (Sequence Retrieval System) is an information indexing and retrieval system presently designed for libraries with a flat file format. It allows a user simultaneous and interactive searching from a menu for character- string (keyword) information in several networked and implicitly or explicitly cross-referenced libraries. Presently about 30 different flat file and indexed libraries (e.g. SWISSPROT, EMBL, PDB) can be indexed and searched. SRS is already installed on a number of EMBnet nodes; namely, Norway, the Netherlands, Switzerland and Spain. It is

proposed, with the help of a computer programmer to be paid by funds requested in this grant application, to extend the retrieval scope of SRS to data files of relational and object-oriented systems (e.g. TFD or Medline) and to allow indexing for retrieval of numerical values and information according to numeric ranges (e.g. SWISSPROT sequences with residue lengths 300 to 400).

The proposal also includes a workshop to be held at EMBL for EMBnet managers after completion of the extensions. The intent is training for the usage and maintenance of the SRS package as well teaching the writing of format specifications for new databases to be added to the library network.

SRS is running well and performing excellently for EMBnet and non- EMBnet users.

F. WWW/XMOSAIC Project

With the advent of Xmosaic the popularity of WWW (World Wide Web) for providing information over the network has increased dramatically.

Xmosaic is a hypertext WWW/browser which allows the presentation of text, (many different fonts) graphics, (gif and mpeg) and audio (au aiff) so that a true multimedia presentation can be made over Internet. WWW has interfaces to gopher, wais, ftp, finger, X500, archie and Whois and as such represents an ideal method of distributing, information and data.

If in the future EMBnet will be considering the distribution of data other than sequence data then it would be advantageous to gain experience with WWW/ Xmosaic because of the excellent typographic capabilities that are offered. At the moment there is no WWW server in Europe which is devoted to molecular biology, and as a pilot project the Center from Scientific Computing would like to develop a EMBnet WWW server. The pilot project would have the following aims.

Develop a WWW home page for EMBnet which would include:

(a) Hypertext document of the EMBnet brochure.
(b) Logos of all the EMBnet nodes.
(c) Photos of all the EMBnet managers.
(d) Links to all the EMBnet gophers.
(e) Links to all the EMBnet ftp sites.
(f) Links to relevant wais sources in molecular biology.

To date there has been some developmental work in converting the brochure ascii file obtained from the UK, first into LaTex format and then using LaTex2html to produce a brochure html hypertext file which can be read by Xmosaic. The URL for this document is appended below. This rudimentary home page will be demonstrated at the next EMBnet business meeting with links to the EMBNet nodes in Finland, Norway and Holland, which should give an outline of the power and functionality.

All nodes are now entered on the service which gives an 'up-to- the- minute' image of EMBnet; and serves for users to gain the most immediate information on the services on offer.

G. NDT protocol on nodes not yet included in network

The NDT protocol-based client/server application has been successfully installed on twentyfive nodes of which 12 are EMBnet national nodes (Austria, Belgium, Denmark, EMBL, Finland, Greece(2), Italy, the Netherlands, Norway(3), Spain(2), Sweden(10), Switzerland, United Kingdom (SeqNet-Daresbury)), and the rest are national subsidiary nodes.

In the time between expiration of the BRIDGE contract and the writing of this report, three more nodes have been included, one national (Israel), and two subsidiary nodes (Spain, Sweden).

Some of the national nodes collaborate in a lattice fashion to the purpose of transporting data as fast as possible by any available route to the end nodes (crossover points). At the time of writing, nine nodes have established such crossover activity. To establish a pan-European EMBnet data distribution using the NDT protocol the following nodes should be included in the lattice: France (Bisance & Genethon), Germany (DKFZ & MIPS), Italy (CNR-Bari), Portugal, Switzerland (Roche), UK (HGMP-RC), and Hungary.

Some of the remaining nodes have excellent network connection, making an installation over the network possible. Other nodes are suffering from substantial delays in data transfers which makes it a very tedious task to ensure a proper installation (and of course those nodes would be the ones that would benefit the most).

NDT is regarded as one of the major successes of the EMBnet programme.

H. EMBL Keyword Structuring

The EMBL data library sequence file is a core feature of EMBnet's activities. The usefulness of any database is related to its retrieval system. One of the widest used search criteria is the 'keyword'.

The EMBL Keyword Index is an alphabetic list. Unfortunately it is inconsistent and has many redundancies; it also lacks a rational economy. This project has taken the index and structured the keywords into keys: DNA, RNA and Protein. Computer programmes for managing the data structuring have also been produced. These form the basis of a user- friendly interface being developed.

Some 15000 keywords have been checked. But this is an on-going programme given the constant evolution of the databases in question. A window interface is also under development to retrieve EMBL entries via the keywords and vice versa.

J. Bioaccellerator

DNA and Protein databases and sequence analysis programs have become critical tools in biological, medical and genetics research. The load on INN's DAPSAS (DNA and Protein Sequence Analysis Service) at the Weizmann Institute of Science computer in the Department of Biological Services for frequently requested services like Database Searches, FASTA, TFASTA, Profilesearch and sequence multi-alignment is rapidly increasing with the ever expanding size of the databases. This is a situation most likely occurring at almost every EMBnet site. This proposal allows All EMBnet nodes to have free access to the Weissman service for a 2 month evaluation period to check out the possibilities of such a service for his own Database Search needs. It should be viewed as a pilot service for the possible installation of one or more dedicated service facilities for all EMBnet nodes for these frequently requested services, by installing a network-wide special purpose computer system, the BIOCCELERATOR.

K. Porting HASSLE to the PC environment

The introduction of HASSLE to the biocomputing environment has brought many advantages for EMBnet node managers as well as for users. In particular the sharing of resources, both in terms of CPU and the databases installed at the HASSLE servers, are very beneficial to all. The current implementation of HASSLE permits users on the EMBnet to use remote CPU and databases without needing to learn

to use another system or another application. The main advantage here being that HASSLE is running in the background without producing interference.

A natural step in this development is to bring the advantages of this system closer to the users. The main idea of this project thus being to promote the use of the EMBnet nodes from the desktop. Using the Windows API (Application Programmers Interface) we wish to develop a set of HASSLE-based applications that would run under the Microsoft Windows environment which will provide access to programs such as Fasta, Blast, Fetch, Medretrieve,etc. Most importantly, the use of HASSLE allows the end user to work on the network without the need to specify a server system. In contrast to WWW, GOPHER, Entrez which need a configuration or make use of meta- indices or special index-pages the HASSLE resource discovery allows to utilize multiple sources.

L. Prototype Volume Database of Three-dimensional Macromolecular Structures

This project examines the problems of information access arising in the field of three-dimensional structural determination by means of processing images through several types of microscopy. The result will be a distributed databank containing information obtained by computer processing of image data from the transmission electron microscope. The databank will be distributed through WWW.

MAJOR SCIENTIFIC BREAKTHROUGHS

Among the most important scientific achievements, noteworthy are the new protocols designed for the synchronisation of data transfer, e.g. the Network Data Transfer Protocol (NDT), retrieval systems, e.g. the Sequence Retrieval System (SRS) and a software to manage distributed processing on heterogenous environment, HASSLE. A project is still underway for the improvement of EMBL database keyword structure, which will ease retrieval of data and updating procedures. The EMBnet User Interface Menu System and the establishment of a central service (the Biocellerator) for all EMBnet nodes to access and test were first steps in what is expected to be a continuing plan to make even better use of national resources at a European level.

MAJOR COOPERATIVE LINKS

EMBnet entered this three year grant period as a collection of bioinformatics nodes organised under a cooperative agreement to execute R&D and training activities. The group ends the period with a formal legal structure, an expanding membership, clear rules for the organisation of its research and other activities, and a vigorous plan for the future. It can also claim an impressive series of training and R&D successes.

To ensure that EMBnet's activities and expertises are better known, EMBnet has produced a brochure which is also available in a continually updated form as an ftp file. This brochure shows clearly the size and range of EMBnet, its tasks and expertises and its increasingly vital role in serving Europe

The EMBnet News file on EMBnet has been used to promote EMBnet activities and detail items of general interest. This is now being used as a central source for national nodes to remove relevant materials for use in their own promotional and educational materials. This is seen as an on-going programme and will be expanded in further programmes, in association with promotional/educational 'Road Shows' for industry — advising them of the advantages and services available through bioinformatics.

The Annual EMBnet meetings were initially of a business and administrative nature although they provided, and provide, a key opportunity for staff to exchange ideas, comment on common problems and devise composite solutions/ projects for identified problems and needs. Each meeting therefore had a scientific component and this was expanded in 1993 into a pure science symposium. With the support of EMBnet's industrial node, Roche, a one day symposium on Protein and Genomic Analyses in the Information Age was held for EMBnet members and the scientific general public. This was a great success and illustrated the interest in such high level meetings as well as pushing the network to the fore. Plans for a 1994 Workshop and a 1995 meeting are in progress.

The first EMBnet trials took place among 5 nodes in 1988. Fourteen centres (nodes) agreed to the decision to form EMBnet, in May 1989 and only twelve of them could sign the BRIDGE contract, namely Israel and Hoffman La Roche did not qualify, the former for not being part of the EU and the latter for being industrial). In this grant period, EMBnet membership has greatly grown and the original nodes now have added about ten new nodes, at present EMBnet counts 17 national nodes and 7 special purpose nodes. EMBnet has always paid particular attention to supporting the applications as start-ups from Southern Mediterranean and Eastern European countries. Right now, 4 additional nodes, Turkey, Poland, Canada and Ireland, have started discussions as to joining as full or associate members.

PUBLICATIONS

EMBnet has enabled a great deal of high quality molecular biology research to be carried out, and has been the source of a large number of research articles covering various aspects of bioinformatics research and development.
The following pages give an impression of the kind of work carried out in the recent past.

Joint publications

Bleasby, A.J., Griffiths, P., SERC Daresbury Laboratory; Harper, R., Finnish State Computer Centre, Tietotie; Hines, D.P., Hoover, K., IntelliGenetics, California; Marshall, S.E., O'Reilly, N., University College Dublin; and Sundvall, M., Uppsala University (1992). Electronic communications and the new biology. Nucleic Acids Research, **20**, 4127- 28

Hegyi, H. and Pongor, S.(1992). Predicting potential domain-homologies from FAST A search results. Computer Applications in the Biosciences, in press Kenter, M., Otting, N., Anholts, J., Leunissen, J., Jonker M. and Bontrop E. (1992). Evolutionary relationships among the primate Mhc-DQA1 and DQA2 alleles. Immunogenetics, **36**, 71-78.

Rysavy, F.R., Bishop, M.J., Gibbs, G.P., Williams, G.W. (1992). The UK Genome Mapping Project online computing service. Computer Applications in the Biosciences, **8**, 149-154.

Savakis, C. and Doelz, R. (1993). Contamination of cDNA sequencers in databases. Science, **259**, 1677

Savakis C. and Doelz R.. 1993. Contamination of cDNA sequences indatabases. Science **259**: 1677-1678.

Individual publications

Akrigg, D., Attwood, T.K., University of Oxford; Bleasby, A.J., SERC, Daresbury; Findlay, J.B.C., North, A.C.T., Maughan, N.A., Parry-Smith, D.J., University of Leeds; Perkins, D.N. and Wootton, J.C. (1992). SERPENT — an information storage and analysis resource for protein sequences. Computer Applications in the Biosciences, **8**, 295-6

Alard, P. and Wodak, S. (1990). Detection of cavities in a set of interpenetrating spheres. Journal of Computational Chemistry, **12**, 918-922.

Amthauer, R, Kodukula, K, Brink, L, Udenfriend, S. (1992). Phosphatidylinositol- glycan (PI-G)-anchored membrane proteins: requirement of ATP and GTP for translation-independent COOH-terminal processing. Proc. Natl. Acad. Sci. U.S.A , **89 (13)**, 6124-8.

Assinder, S.J., de Marco, P., Sayers, J.R., Shaw, L.E., Winson, M.K. and Williams, P.A. University of Wales, Bangor. (1992). Identical resolvases are encoded by Pseudomonas TOL plasmids pWW53 and pDK1. Nucleic Acids Research, **20**, 5476

Bex, F., Pierard, A., Desmyter, A., Dreze, P., Colet, M. and Couturier, M. (1986). The carboxy-terminal end of the mini-FE protein is essential for E gene repression and mini-F copy number control. J. Mol. Biol., **189**, 293-303.
Bishop, M.J. (1989). Computing. In: Howe, C.J. and Ward, E.S. (eds), Nucleic acids sequencing: a practical approach. Oxford: IRL Press.
Bishop, M.J. (ed.) 1994. Guide to human genome computing. Academic Press, London.
Blommers, M.J.J., Bruins Slot H.J., van der Marel, G.A., van Boom, J.H., Hilbers , C.W. (1990). Conformational analysis of the 3'-5'-cyclic dinucleotide d«pApA» by means of molecular mechanics. J. Biomol. Struct. Dyns. **8**, 523.
Carazo, J.M. (1992). The accuracy of 3D reconstructions from incomplete information and the use of restoration methods. In: J. Frank (ed.), Electron Tomography. Plenum Press, pp. 117-164.
Carazo, J.M., Marco, S., Abella, G., Carrascosa, J.L., Secilla, J.P. and Muyal, M. (1991). Electron microscopy study of GroEL Chaperonin: Different views of the aggregate appear as a function of cell growth temperature. J Struct. Biology, **106**, 211-220.
Carazo J.M., Marabini R., Vaquerizo C. y Frank J. Communication to the congress ICEM 13 (International Congress on Electron Microscopy'), 17-22 July 1994 (Paris, France), 'Towards a Data-Bank of three-Dimensional molecular structure: A WWW based prototype'.
Carrascosa, J.L., Carazo, J.M., Herranz, L., Donate, L.E. and Secilla, J.P. (199 0). Study of two related configurations of the neck of bacteriophage phi29. Computers and Mat hematics, **20 (4)**, 57-65.
Cunningham, P.R, Nurse, K, Bakin, A, Weitzmann, C.J, Pflumm, M, Ofengand, J. Interaction between the two conserved single-stranded regions at the decoding site of small subunit ribosomal RNA is essential for ribosome function. Biochemistry, **31 (48)**, 12012-2 2.
Danho, W, Tilley, J.W, Shiuey, S.J, Kulesha, I, Swistok, J, Makofske, R, Michalewsky, J, Wagner, R, Triscari, J, Nelson, D, et. al. (1992). Structure activity studies of tryptophan30 modified analogs of Ac-CCK-7. Int. J. Pept. Protein Res., **39 (4)**, 337-47.
Debue, M. Gauthier, P. Hackel, C., Van Elsen, A. Herzog, A., Bigaignon, G. and Bollen, A. (1991). Detection of *Borrelia burgdorferi* in biological samples using the polymerase chain reaction. Res. Microbiol., **142**, 565-572.
de Giorgi, C., d'Alessandro, A. and Saccone, C. (1992). Mitochondrial DNA detect ion and copy number determination in the spermatozoa of the sea urchin Arbacia Lixula. Bioch. and Bioph. Res. Communic., **182 (3)**, 1454-1459.
Doelz, R. (1990). The Biocomputing Tutorial. CSC, Helsinki.
Doelz, R. (1992). Computer Networks in the Bioinformatics Infrastructure. Bioinformatics, **1**, 40-46.
Doelz, R. (1992). The EMBnet Project. Computer Networks and ISDN Systems, **25**, 46 4-468.
Doelz, R. (1993). Use of the GCG programs. In: Griffin, A. and Griffin, H.G. (ed s.), Computer Analysis of Sequence Data. 1993. Chapters 5-17. Humana Press New Jersey, in press.
Etlinger, H.M. (1992). Carrier sequence selection — one key to successful vaccines. Immunol. Today, **13 (2)**, 52-5.
Hackel, C., Houard, S., Portaels, F., Van Elsen, A, Herzog, A. and Bollen, A. (1 990). Specific identification of *Mycobacterium leprae* by the polymerase chain reaction. Molecular and Cellular Probes, **4**, 205-210.
Harper, R. (1992). Networking for the novice. Microcomputers in Biochemistry (A practical Approach). Editor C. Bryce, p. 267-296. Oxford University Press 1992.
Hegna, I., Karlstrom, E.S., Lopez, R.,Kristensen, T. and Kolst, A. (1992). A type-III DNA restriction and modification system in *Bacillus cereus*? Gene, **144**, 149-150.
Heumann K., George D.G., and Mewes H.W. A New Concept of Data Distribution for Sequence Data Banks. Comp. Appl. Biosc. (1994) in press
Kolakowski Jr., L.F., Leunissen, J.A.M. and Smith, J.E. (1992). ProSearch: Fast searching of protein sequences with regular expression patterns related to protein structure and function. Biocomputing, **13**, no. 6, 919-921.
Kristensen, T., Lopez, R. and Prydz, H. (1992). An estimate of the sequencing error frequency in the DNA sequence databases. DNA Sequence 1992, 343-346.
Kunsch, C, Ruben, S.M, Rosen, C.A. (1992). Selection of ptimal kappa B/Rel DNA-binding motifs: interaction of both subunits of NF-kappa B with DNA is required for transcriptional activation. Mol. Cell Biol., **12 (10)**, 4412-21.

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992). CpG Islands as Gene Markers in the Human Genome. Genomics, **13**, 1095-1107.
Leunissen, J.A.M., van den Hooven, H.W., de Jong, W.W. (1990). Extreme differences in charge changes during protein evolution. J. Mol. Evol. 31, 33-39.
Linder, P., Doelz, R., Gubler, M., and Bickle, T.A. (1990). An anticodon nuclease gene inserted into a hsd region encoding a type I DNA restriction system. Nucl. Acids Res. **18**, 7170
Lopez, R., Kristensen, T. and Prydz, H. (1992). Database Contamination. Nature, 355, 211.
Lopez R. The Norwegian EMBnet node. UniNytt 4 (1993)
Lopez R. EMBnet NBS-Nytt 3 (1993) 16-17
Lopez R., Larsen F. and Prydz H. Evaluation of the exons predictions of the GRAIL software Genomics (in press)
Maconchie, M.K., Brown, S.D.M. and Greenfield, A.J. St. Mary's Hospital Medical School, London. (1992). Sequence analysis of two exons from the murine dystrophin locus. Mammalian Genome, **2**, 64-8
Marabini R., Vaquerizo C., Fernandez J.J., Carazo J.M., Ladjadj M., Odesanya O. and Frank J. Communication to the congress VMC-94 (Visualization in Biomedical Computing-94, organized by Mayo Foundation), 4-7 October 1994 (Rochesters, USA), 'On a prototype for a new distributed data base of volume data obtained by 3D imaging'.
Narahara, A., Naterstad, K., Kristensen, T., Lopez, R., Bork, P.and Kolst, A. (1992). Cloning of a gene from Bacillus cereus with homology to the mreB gene form *Escherichia coli*. Gene, **122**, 181-185.
Parsons, J.D., Brenner, S. and Bishop, M.J. (1992). Clustering cDNA sequences. Computer Applications in the Biosciences, **8**, 461-466.
Pesole, G., Attimonelli, M., Preparata, G. and Saccone, C. (1992). A statistical method for detecting regions with different evolutionary dynamics in multialigned sequences. Molecular Philogenetics and Evolution, **1 (2)**, 91-96.
Pesole, G., Prunella, N., Liuni, S., Attimonelli, M. and Saccone, C. (1992). Wordup: An efficient algorithm for discovering biologically significant patterns in DNA sequences. Nucleic Acid Research, **20**, 2871-2875.
Pesole, G., Sbis, E., Mignotte, F. and Saccone, C. (1991). The branching order of mammals: Phylogenetic trees inferred from nuclear and mitochondrial data. J. Mol. Evol., **33**, 537-542.
Pizzi, E., Attimonelli, M., Liuni, S., Frontali, C. and Saccone, C. (1991). A simple method for global sequence comparison. Nucleic Acids Res., **20**, 131-136.
Polner, G., Skerl, V. and Pongor, S. (1992). Detection of protein sequence patterns in database search results. Prot. Seq. Data Anal., in press.
Pongor, S. (1992). Bioinformatika Biokemia, **16**, 145-153
Ritter, O. et al. (1992). Risc of cystic fibrosis based on marker genotype data. In: Macek, M., Ferguson-Smith, M.A. and Spala, M. (eds.), Early fetal diagnosis: Recent progress and public health implications. Charles University Press.
Rouch, D., F. Wright University of Birmingham; and Bleasby,.A.J. SERC, Daresbury; (1992). Improving front-line user support: proposals for greater efficiency in computational molecular biology. BINARY, 4, 29-32
Rysavy, F.R. (1992). Users and services providers: Interoperability; Experience from the UK Human Genome Mapping Project Users Group. Proceedings of the 3rd Joint European Networking Conference, Innsbruck, Austria, May 11-14, 1992. Amsterdam: Elsevier, in press.
Siezen, R.J., de Vos, W.M., Leunissen, J.A.M., Dijkstra, B.W. (1991). Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteases. Protein Engineering, 7, 719-737.
Simon, G. and Pongor, S.(1992). ICGEBnet: The UNIDO computer resource for molecular biology. Bioinformatics, 1,12
Suhai, S. (1989). Computational methods in Cancer research. Interdisc. Science Reviews, **14**, 225-232.
Suhai, S. (1990). Computer-aided analysis of biomolecular sequence and structure in genome research. In: Geldermann, H. and Ellendorf, F. (eds.), Genome analysis. VCH Publishers, New York, 1990, pp. 81-104.
Suhai, S. (1990). Modelling of protein structures on the basis of sequence data. In: Tschesche, H. (ed.), Modern methods in protein- and nucleic acids research. W. de Gruyter, Berlin, 395-422.

Suhai, S. (1991). The electronic structure of parallel-beta-pleated sheets in proteins: Ab initio computation including electron correlation. Int. J. Quant. Chem., **40**, 559-576.
Thieffry, D., Verdin, E. and Colet, M. (1992). Regulation of HIV expression: a logical analysis. Lectures in Biomathematics, in press.
Valpuesta, J.M., Fujisawa, H., Marco, S., Carazo, J.M. and Carrascosa, J.L. (1992). Three-dimensional structure of T3 connector purified from over-expressing bacteria. J. Mol. Biol., **224**, 103-112.
Veljkovic,V., Metlas, R., Raspopovic R. and Pongor S. (1992): Spectral and sequence similarity between vasoactive intestinal peptide and the second conserved region of human immunodeficiency virus type I. envelope glycoprotein. Biochem. Biophys. Res. Comm., **189**, 705-710.
Williams, G., Gibbs, G.P. (1990). Automatic updating of the EMBL database via EMBnet. Computer Applications in the Biosciences, **6**, 122-123.
Zapata, E.L., Benavides, J.L., Plata, O.G., Rivera, F.F., and Carazo, J.M. (1990). Image template matching on hypercubes SIMD computers. Signal Processing, **21**, 49-60.

Integrated data and knowledge base of protein structure and sequence (BIOT CT-910271)

COORDINATOR:
1) S.J. WODAK, ULB, Bruxelles, B
PARTICIPANTS:
2) C. SANDER, EMBL, Heidelberg, D
3) J. THORNTON, Univ. College London, London, GB
4) P. GRAY, Univ. Aberdeen, Aberdeen, GB
5) C. RAWLINGS, ICRF, London, GB
6) M. MOREAU, BIM, Everberg, B

BACKGROUND INFORMATION

A number of computer programs, developed outside this project, were made available to the project in whole or in part; some of these were further developed within the project. These programs comprise: the databases SESAM-ALI *(partner 1)*, IDITIS *(partner 3)*, TOPOL and the supporting system PAPAIN *(partner 5)*, P/FDM and AMAZE *(partner 4)*. AMAZE uses a 3D presentation of the object classes to help users formulate queries for applications related to antibody design. The antibody data is taken from a separate database funded by SERC.

OBJECTIVES AND PRIMARY APPROACHES

One of the main objectives of this project has been to develop a computer environment for protein structure through integration of data, algorithms and knowledge based computational approaches, achieved by the use of a common design schema. The primary approaches included:

1. Design of a standard schema for manipulating and exchanging molecular biological data entities among partners and between software developed by them
2. Cross validation of data and software developed by different partners in the context of protein structure prediction and modelling
3. Shared development of user interfaces (front-end) and efficient data storage programs (back-end)
4. Integration of advanced methods of database management and AI technology into the field of molecular biology
5. Improved knowledge based methods for protein structure prediction.

RESULTS AND DISCUSSION

A. All partners

1. Definition of the standard schema

All partners took an active part in meetings and discussions on the design of a consensus conceptual schema for the representation of protein structure and sequence data items and the relations between them. Consensus was reached concerning a core schema representing primarily the raw protein structure data. This consensus schema has now been implemented by various partners, and served as a guide for specifying the data structure for the new CIF (Crystallographic Information File)

standard, developed by a working group designated by the IUCr (International Union of Crystallography).

A Schema Definition Language (SDL) for interchange of project database designs was developed in a close collaboration between *partner 1* and *partner 5*. The SDL was designed to facilitate the exchange of designs for the BRIDGE schema based on object-oriented semantic data model. It is an abstract schema definition language for describing the attributes of each entity and its relations with other entities. In particular, the entity class, primary key, simple attributes, and relations are defined separately using a block structuring convention. The SDL is a very valuable tool, which we hope will be more extensively exploited in the future.

B. Partner 1

1. Enhancement of SESAM and its user interface ALI

Ways of representing data on multiple sequence alignments were devised and such data were entered for 28 protein families including the globins, cytochromes, lysozymes, and microbial RNases. An exhaustive repertoire of sequence patterns associated with common secondary structures, as well as their respective territories, were stored in SESAM.

Commands were developed which are capable of interactively updating the metaknowledge in ALI about the core database. The procedures that build automatically complex SQL queries have been improved and a tree structure has been introduced into the ALI menu. SESAM-ALI have been ported to several hardware platforms, including the IBM compatible PC platform running UNIX.

2. Automatic procedures for comparison and classification of protein structure motifs

Two different procedures were developed. One, automatically identifies families of turn motifs adopted by contiguous polypeptide segments of 10 to 30 residues long in the database of known protein structures. It does so by clustering overlapping segments with similar conformations from all the protein structures based on the rms deviations of backbone atoms, and by applying filters that compare backbone dihedral angle values.

A second procedure automatically derives the optimal alignment of the 3D structures of two proteins, using as sole similarity measure the rms deviations of the backbone atoms after coordinate superposition. The algorithm reliably provides optimal alignments for distantly related proteins, and can be used to analyze rigid body movements, as well as very local structural deformations. Its excellent performance in defining the common structural core in distant homologs and in identifying hitherto undetected similarities between substructures was demonstrated.

3. Analysis of sequence-structure relationships and their use in structure prediction programs

Different types of effective potentials were derived from the database of known structures based on statistical analyses of the relationships between the amino acid sequence and different descriptions of the protein structures. These potentials have been used to evaluate sequence-structure compatibility in threading procedures, and in structure prediction. Algorithms to predict well defined conformations in short peptides and to identify protein segments with well defined conformational preferences were developed and shown to yield results in excellent agreement with experimental data.

C. Partner 2

1. Database information on sequence families

The HSSP database of protein families contains a set of biological relations between protein structures and all homologous sequences, and serves as an important link between the structure databank (PDB) and the protein sequence databank (SwissProt). The secondary structure information of known protein structure is now extracted routinely from HSSP and reported in the corresponding SwissProt entry. Homology pointers from a protein sequence in SwissProt to a known structure will be included shortly. Using HSSP, sequence discrepancies between the entries in the PDB and SwissProt are reported in an automatic fashion. The HSSP-files also serve as input for a newly developed secondary structure prediction method using a neural network. The prediction method rates a sustained 72% average accuracy for water-soluble globular proteins, for cases in which at least two homologous sequences are known. An electronic mail service for secondary structure prediction *(predictprotein@embl-heidelberg.de)* was installed and is used heavily (appr. 1000 requests per month).

2. Efficient database production on parallel computers

The program generating the HSSP-database has been implemented on a range of parallel multiple-instruction-multiple-data (MIMD) computers, with the aim of reducing the time to update the HSSP-database from weeks to a few days. Production runs begun early in 1994.

3. Data field extraction from sequence and structure databases

PickProt is a set of scripts based on the Perl language, performing data extraction from protein structure databases (PDB, DSSP, HSSP). It can also be used to enumerate, select and tabulate data entities. Different sets of constraints result in the generation of various lists of protein chains that can be used for further analysis.

4. User interface for work with heterogeneous databases

ProtQuiz and GeneQuiz are new X-window based programs that allow browsing through molecular flat-files of different origin (PDB, DSSP, HSSP, FSSP, SwissProt). They provide a flexible user interface and can be used as a visualization tool for molecular structures and sequence information. The software is object-oriented and was written in C++, using the ET++ application framework (ETH, Zürich). The actual implementation of internal data structures reflects the proposed conceptual scheme developed in the present project.

5. Search tools for protein structure databases

A new type of database search facility was designed in the package WHATIF, which allows queries about residues and their sequence (1D) and structure (3D) environments in an efficient way. The tools complement those provided by the Bruxelles *(partner 1)*, London *(partner 3)* and Aberdeen *(partner 4)* groups.

6. Network access to databases and programs

Access to the HSSP and related databases (DSSP, FSSP) and to the ProtQuiz program is now provided via international electronic network. On Internet, files can be downloaded using ftp (file transfer protocol) from *ftp.embl-heidelberg.de* (directories /pub/databases/protein—extras and /pub/software/unix/protquiz) and

electronic mail access is provided by the server *netserv@embl-heidelberg.de* (send the message text: help).

D. Partner 3

With reference to our objectives:

1. We have participated in numerous discussions on the schema design, leading to the final model.
2. We have extended the methods for structural comparison, fold classification and motif recognition [1-11].
3. We have produced an automated procedure for clustering the protein structure database [1] and analysed in detail the alpha plus beta fold family [2]. The classification scheme was made available to participants prior to publication for incorporation into their databases. This work is continuing towards developing consensus templates for the different fold families, for classification and to aid fold recognition.

E. Partner 4

Our initial objective was to represent entities of biological interest as objects having relationships to other objects, all stored in an object-oriented database according to a schema agreed by the consortium. Instead of using a commercial OODB we used the P/FDM software developed on previous projects. This allows a variety of storage modules, and we had to develop a way to use a SYBASE relational database to store objects for which our system had assigned object identifiers. This assignment is a crucial step, and it overcomes the problems of ambiguity that arise in flat CIFiles through the lack of a clear way to uniquely identify all objects.
The main advantage of storing objects in this fashion is that we can then apply computational methods to them in order to calculate interesting derived properties and relationships. These methods are stored in the database as pieces of code with type descriptors, and they behave like convenient small tools which fit the objects. They can even be fitted together in order to compute complex queries as they arise, through the P/FDM query interface on a workstation. This saves one from working with low-level SQL on relations, and then having to convert results for input to an application package, and so on.

We have fulfilled our objective by populating a database with 137 representative 3-D protein structures identified by Orengo and Thornton (1993), together with their fold families and related families (as objects). We have also integrated sequence data for related proteins of unknown structure in the case of antibody families. Our query interface can be used to search and calculate with all this data and attached methods, and it is designed to make efficient use of the speed of relational storage.

F. Partner 5

1. Addition & Validation of data and knowledge

a. Representing Protein Structure

Several geometries used to describe packed protein secondary structure element pairs abstracted as pairs of vectors were compared to determine the best geometric features for describing both packed and non-packed elements in proteins. This

description of three-dimensional relationships has been applied in the area of protein structure comparison and will be of value for many other applications.

b. Validation of Protein Topology Data

A number of programs were developed using the logic programming language Prolog to check the consistency and completeness of the topological description of protein structure used in the TOPOL database. These programs are now integrated into the PAPAIN (Protein Analysis and Prediction using Advanced INformatics) system.

c. Representing Multiple Labellings on Sequence Data

A general method for representing alternative co-existing interpretations of protein sequence data was developed. This was achieved by separating the concept of a physical polypeptide chain segment from a logical segment that could be composed of one or more physical segments or parts thereof. This also helps to resolve the problem of residue labellings on discontinuous regions of the polypeptide chain.

2. Software and User Interfaces

Efforts were focussed on extending the data and features within our prototype software called PAPAIN, which contains both a range of conventional sequence analysis procedures and more advanced features such as constraint-based protein topology prediction.
The developed features comprise:

(1) New graphical interfaces for predicting protein topology and extensions to the protein topology diagramming tool enabling alpha/beta topologies to be displayed.
(2) An interface to the PROSITE database and the PROSEARCH programs for locating the position of sequence motifs.

3. Integrating Databases

Programs to select subsets of data from the database IDITIS *(partner 3)* and to convert them into Prolog clausal form were developed. Considerable work was required to convert the IDITIS data to Prolog form so that the PAPAIN system can be underpinned by a more reliable and well-validated database. The opportunity was taken to use the fact that IDITIS provides alternative secondary structure assignments (the original author assignments as reported in the PDB, and those derived by the Kabsch and Sander hydrogen bonding rules). Means for viewing alternative secondary structure assignments have been developed and are now exploited in PAPAIN.

4. Linking Protein Structure to Gene Structure

A database recording the exon fingerprints (FINEX) was developed: it records, for each gene, the number of exons, their size and the phase of the intron/exon junctions. This database has proven to be useful for a rapid identification of related genomic sequences. In the future, links will be developed with the structure databases to allow analysis of relationships between protein and gene structure.

5. Structure comparison

In collaboration with *partner 3* and William Taylor (MRC NIMR), a comprehensive review of protein structure comparison methods was written [1]. It develops a general nomenclature for categorising the wide diversity of structure comparison

methods, that, unlike earlier reviews on the topic, unifies apparently disparate techniques.

G. Partner 6

This partner closely collaborated with *partner 1* in:

(1) the design and implementation of the SESAM database schema for representing information on multiple sequence alignments,
(2) the development of procedures in ALI for accessing and viewing sequence mapped information, including motifs of secondary structure arrangements,
(3) the optimization of the automatic query capacity of ALI,
(4) the development of procedures based on expert problem solving approaches for the programs that load data from the protein databank (PDB) into the SESAM database, with the effect of significantly reducing human intervention in the loading process.

This partner also provided continuing guidance on the internals and optimization of SYBASE, the commercial RDBM system used to implement SESAM.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

Partner 1

The structure prediction methods based on the effective potentials have found many applications, some of which in the pharmaceutical industry. In particular the programs PRELUDE and FUGUE, which respectively predict well defined conformations of short peptides, and identify locally structured regions in complete proteins were installed in several companies and research centers in Europe and the US. The relational protein sequence and structure database SESAM has also been installed in several locations in Europe and the US. It is also accessible for interrogation over the electronic network, by scientists world-wide.

Partner 2

The database HSSP containing information on protein structures and all homologous sequences was developed, and links were established between it and the protein structure (PDB) and protein sequence (SwissProt) databanks. Based on HSSP the secondary structure prediction programs was raised to the level of 72% on the average.

Partner 3

The clustering of motifs from the Protein Databank has provided an overview of the 'world' of protein structures, and led to suggest the occurrence of superfolds and singlet structures [11]. The methods to classify the structures will in the future be used to provide input data for 'IDITIS' — a relational database of protein structures, which is commercialised by 'Oxford Molecular'.

Partner 4

Integration of sequence and structure data in an object-oriented framework by *partner 4* is the first such implementation using a realistic amount of protein data. The graphic interface (AMAZE) is used by this partner to locate objects via interactive graphics remotely.

Partner 5

The FINEX database is a novel development from the project. Its contribution to the understanding of the evolution of genomic and protein structure and potential applications in protein engineering or the human genome project are being evaluated.

Partner 5 has achieved an integration between the logic based system PAPAIN and IDITIS the relational database developed at UCL *(partner 3)*

MAJOR COOPERATIVE LINKS

General comments

The present project has led to the development of the consortium for 'Data Validation' supported by the EC Biotechnology program. *Partners 1, 2, and 3* have furthermore developed close links (as alpha-collaborators) with US laboratories, and are currently involved in planning future developments with them for protein data validation, deposition, archiving and retrieval. This will be done in collaboration with the European Bioinformatics Institute, and with different collaborative networks hopefully funded by the 4th framework programme of the EU.

Specific collaborations

Partner 1

With *all partners* on the development of the standard conceptual schema.
With *partner 3* (UCL) on comparing structure alignment methods
With *partner 4* (Aberdeen) on the use of SYBASE
With *partner 5* (ICRF) on the development of the schema definition language (SDL).
With partner 6 (BIM) on database design problems, user interfaces, and load procedures

Partner 3

With *partners 4* and *5* on integrating the database IDITIS and data thereof with the object oriented, and logical databases at Aberdeen and ICRF respectively.

Partner 4

Worked closely on the use of SYBASE with *partner 1.*
Worked closely with *partner 2* to develop external schemas (or views) for our object database. It is now clear that different users need very different views (or presentations) of the one shared underlying set of objects, and that further development along the lines we have established would enable much wider use of the database.

Partner 5

Links to UCMB *(partner 1)* made an important contribution in the development of the Schema Definition Language.
Links to UCL *(partner 3)* and through them to Oxford Molecular enabled ICRF to use the IDITIS database and has made a substantial contribution to the further development of the PAPAIN system and supporting databases.

PUBLICATIONS

Individual publications

Partner 1

1. Rooman M.J., Kocher J.-P.A. and Wodak S.J. (1991) Prediction of protein backbone conformation based on 7 structure assignments: I. Influence of local interactions. J. Mol. Biol. **221**, 961-979
2. Rooman M.J., Kocher J.-P. and Wodak S.J. (1992) Extracting information on folding from the amino acid sequence: accurate predictions for protein regions with stable conformation in absence of tertiary interactions. Biochemistry, **31**, 10226-10238
3. Rooman M.J. and Wodak S.J. (1992) Extracting information on folding from the amino acid sequence: role of consensus stable regions in homologous proteins. Biochemistry, **31**, 10239-10249
4. Rooman, M.J., Wodak, S.J. and Kocher, J.-P. (1993) Prediction of regions with well defined conformational preferences in proteins, and their relevance to protein folding. In M. Doyama, J. Kihara, M. Tanaka and R. Yamamoto (eds.) *Computer Aided Innovation of New Materials II.* Elsevier Science Publishers b.v., 1129-1235.
5. Wodak S.J. and Rooman M.J. (1993) Generating and testing protein folds. Current Opinion in Structural Biology, **3**, 247-259
6. Kocher J.-P.A., Rooman M.J. and Wodak S.J. (1994). Factors influencing the ability of knowledge based potentials to identify native sequence-structure matches. J.Mol.Biol. **235(5)**, 1598-1613
7. Rooman M.J., Kocher J.-P.A., Wintjens R. and Wodak S.J. (1994) Knowledge based potentials for predicting the 3-dimensional conformation of proteins. in *Statistical mechanics, protein structure, and protein substrate interactions*, S. Doniach (ed.,) Proceedings of a NATO meeting held at Cargèse, France, 1993 (in press)
8. Wintjens R.T., Rooman M.J. and Wodak S.J. (1994) Identification of short turn motifs in proteins using sequence and structure fingerprints. Israel Journal of Chemistry **34(2)**, 257-269

Partner 2

1. U. Hobohm and C. Sander (1994). Enlarged representative set of protein structures. Protein Science **3**, 522-524
2. L.Holm and C. Sander (1994). Structural similarity of plant chitinase and lysozymes from animals and phage: an evolutionary connection. Febs Letters **340**, 129-132
3. L.Holm, A. Murzin and C. Sander (1994). Three sisters, different names: 3α, 20b-hydroxysteroid dehydrogenase, dihydropteridine reductase, UDP-galactose 4-epimerase. Nature Structural Biology **1**, 146-147
4. G. Vriend, P.F. Stouten and C. Sander (1994). A novel search method for protein sequence-structure relationships using property profiles. Protein Engineering **7(1)**, 23-29
5. R.Schneider and C. Sander (1993). The HSSP data base of protein sequence-structure alignments. Nucleic Acids Research **21**, 3105-3109
6. L.Holm and C. Sander (1993). Structural alignment of globins, phycocyanins and colicin A. Febs Letters **315**, 301-306
7. L.Holm and C. Sander (1993). Globin fold in a bacterial toxin. Nature **361**, 309-309
8. P. Bork, A. Valencia and C. Sander (1993). Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase and galactokinase families of sugar kinases. Protein Science **2**, 31-40
9. A. Valencia, K.C.Holmes and C. Sander (1993). A new ATP binding fold. A common structure for actin, hexokinase and the Nt domain of Hsc70. Trends in Cell Biology **3**, 53-59
10. L. Holm and C. Sander (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. **233**, 123-138
11. L.Holm, C. Ouzounis, C. Sander, G. Tuparev and G. Vriend (1992). A database of protein structure families with common folding motifs. Protein Science **1**, 1691-1698
12. U. Hobohm, C. Sander, M. Scharf and R. Schneider (1992). Selection of representative protein data sets. Protein Science **1**, 409-417

Partner 3

1. Orengo C.A., Flores T.P., Taylor W.R. and Thornton J.M. (1993). Identification and Classification of Protein Fold Families. Protein Eng. **6(5)** 485-500

2. Orengo C.A. and Thornton J.M. (1993). Alpha Plus Beta Folds Revisited: Some Favoured Motifs. Structure **1**(2) 105-120
3. Thornton J.M. and Gardner S.P. (1993). Database Analysis for Protein Engineering. Chemical Design Automation News 18-21
4. Orengo C.A., Flores T.P., Jones D.T., Taylor W.R. and Thornton J.M. (1993). Recurring Structural Motifs in Proteins with Different Functions. Current Biology **3**, 131-139
5. Orengo C.A. and Thornton J.M. (1993). Protein Fold Families and Structural Motifs. Proceedings of the NATO ARW Workshop on Statistical Mechanics: Protein Structure and Protein Substrate Interactions. (in press)
6. Orengo C.A. and Taylor W.R. (1993). A Local Alignment Method for Protein Structure Motifs. J. Mol. Biol. **233**, 488-497
7. Flores T.P., Orengo C.A., Moss D.M. and Thornton J.M. (1993). Comparison of Conformational Characteristics in Structurally Similar Protein Pairs. Protein Science **2**, 1811-1826
8. Jones D.M., Orengo C.A. and Thornton J.M. (1994). Protein Folds and Their Recognition from Sequence. In M.J.E. Sternberg (ed.) *Prediction of Protein Structure: A Practical Approach.* Oxford University Press
9. Orengo C.A. (1994) Classification of Protein Fold Families. Current Opinions in Structural Biology, **4**, 429-440
10. Taylor W.R., Flores T.P. and Orengo C.A. (1994). Multiple Protein Structure Alignment. Protein Science (in press)
11. Orengo C.A., Jones D.T. and Thornton J.M. (1994). The Distribution of Protein Domain Structures (submitted)

Partner 4
1. Kemp G.J.L., Iriarte J.J. and Gray P.M.D. (1994). Efficient Access to FDM Objects stored in a Relational Database. In D. Bowers (ed.) *Proc. BNCOD-12: 12th British National Conference on Databases.* Springer-Verlag
2. Kemp G.J.L., Jiao Z., Gray P.M.D. and Fothergill J.E. (1994) Combining Computation and Database Access in Biomolecular Computing. In T.Risch & W.Litwin (eds.) *Proc. Int'l Confce on Applications of Databases — ADB'94 (Sweden).* Springer-Verlag

Partner 5
1. Brown N.P., Orengo C.A. and Taylor W.R. (1994). A Protein Structure Comparison Methodology. Eur. J. Biophys. (submitted)
2. Clark D.A. Shirazi J., Rawlings C.J. (1991). Protein topology prediction through constraint-based search and the evaluation of topological folding rules. Protein Engineering **4**, 751-750

Immunoclone and hybridoma database network for Europe (BIOT CT-910257)

COORDINATOR:
L. RÉCHAUSSAT, CERDIC, Nice-Sophia-Antipolis, F
PARTICIPANTS:
Associate contractors:
A. DOYLE, ECACC, Porton Down, UK
T. RUZZON, IST, Genova, I
P. de BAETSELIER, VIB/IMB, Brussels, B
H. KURZVELLY, DIMDI, Köln, D
Subcontractor:
H. LUCIUS, Max-Delbrück Centrum (ex-ZIMB), Berlin, D

BACKGROUND INFORMATION

Immunoclones can be defined as any permanent cell line, obtained by hybridisation, virus transfection or DNA transfer producing by secretion or membrane presentation homogenous substances of immunological interest (monoclonal antibodies, T-cell receptors, interleukins, macrophage factors, etc.). Immunoclones produced in private or public research raise serious problems of identification and availability due to the recent and fantastic expansion of those biological products. Cell line resource banks represent a very little portion of available immunoclones and monoclonal antibodies. Nevertheless scientists and industrials need clear descriptions of available immunoclones for research or development purpose.

OBJECTIVES AND PRIMARY APPROACHES

1. Development of a scientific and technical database on immunoclone descriptions (Immunoclone Database-ICDB) coming from:
— Scientific literature currently published,
— Patent applications to the European Patent Office,
— Industrial and commercial catalogues,
— Participants collections of hybridomas.
Altogether, the target was an input capacity of 2,000 descriptions per month after 18 months.

2. Implementation of a stable consortium agreement, as a model for biological database network in Europe.

3. Development of an end-user interface for improvement of the use of databases for biotechnology in Europe.

RESULTS AND DISCUSSION

A. Major problems encountered

1. Concerning the whole project:

One of the main problems encountered was the lack of electronic communication facilities in involved laboratories. Instead of a simple unique system, the coordinator was obliged to implement specific procedures adapted to each local situation.

The development of automatic conversion software of ASCII format input into ORACLE format and vice versa took more time than expected.

The implementation of data input started later than expected, due to the delay of the EEC funds shipment, which obliged the coordinator to postpone the initial training of the technical staff of participating centres.

2. Concerning centres:

At the local level, some difficulties were raised by the administration for microcomputer acquisition in due time and by some coders who left their position soon after the training sessions.

One centre decided to leave the project, some months after it started.

3. CERDIC reviewed the communication problems and payed a visit to production centres in order to implement an appropriate solution for each case.

B. Database management implementation

1. The *data input and validation software*, was despatched 6 months after project has started and improved continuously in order to integrate new functions related to the *ICDB Vocabulary list and the EMBASE Thesaurus viewing* (with the kind authorisation of Elsevier Science Publishers). A help system and various extended validation routines (262,000 bytes in 20 C- language modules) were added. Five build-in databases are used when the software operates (ICDB Vocabulary, EMTREE Subset, Rules file, Help file and Numbering data).
2. The *main database of ICDB is managed by ORACLE procedures and specific software* (136,700 bytes in 12 modules) for insertion and extraction of records. The internal structure includes 17 information tables and 12 join tables. A preformatted Text table is added in order to facilitate text retrieval.
3. A database was added for the process of journal article references which are extracted from floppy disks containing a selection of bibliographical data, shipped several times per month by CNRS (Pascal), INSERM (Medline) and Elsevier (Embase) . *Automatic Redistribution of Bibliographic Reference Entries software* (ARBRE) was developped in order to read the textfile as delivered by EMBASE, MEDLINE and PASCAL producers and to insert the reference information into the ORACLE database, including the respective database record number. *ARBRE detects the country of origin of the 1st author and assigns each reference to one of the coding centre*, according to a dispatching list. Each list of those references is sent automaticly by electronic mail to the appropriate coding centre.
4. An original software for the automatic process of E-Mail from inputing centres was developped. *Automatic Selecting of Pertinent Input and Reporting software* (ASPIR) reads incoming E-Mail, detects ICDB records of input centres, open the message, start the insertion programme, analyzes the results and reports to the sender and to the database manager. When encountered, the offending records are reproduced with error messages explaining the problem.

C. Hardware and software installation

The initial installation was done during the 1st year of the project. Updates of the database and of the input software were sent and implemented in each centre (except ECACC).

D. Communication procedures

Difficulties encountered during the first period were solved. Appropriate solutions has been found for each centre. The regular use of ASPIR solved most of file transfer difficulties.

E. Data collection

The coordinating centre dispatched among the partners:
— 15,560 bibliographic references,
— 1,160 patents and
— 70 new catalogues from major companies.

F. Data processing implementation and optimisation

Three training sessions were organized at CERDIC for the improvement of input techniques and production rate.

G. User-friendly interface

1. A *menu driven interface* was developped by DIMDI for Immunoclone Database, in order to attract a wider community of scientists to use the online database. For the ICDB user guidance the expert knowledge has been used to *optimize the internal search strategy for antibodies*, either by its reactivity or by authors. Comprehensive description of each search feature of the user guidance has been integrated into *'Help-function'* of the system, in English and in German.
2. A feature specific to ICDB user guidance is the *Hypertext function which links ICDB records with their literature sources* identified in Medline and/or Embase in order to enables the user to perform a simultaneous search.

MAJOR SCIENTIFIC BREAKTHROUGHS

1. Consortium agreement

In addition to the Bridge contract, associate contracts and subcontract, a consortium agreement was finalized in November 1991 and countersigned by each participant. This document states on the rights and duties for input activities, sharing of tasks, copyright, sharing of royalties, local use of the database and distribution policy.

2. Production

20,293 records were produced during the project. In addition, CERDIC processed about 3,900 records on scientific literature for the years 1984 to 1988, for homogenity purpose.

The monthly production met 2,020 records in March 1993, last month of the project.

In total, ICDB on-line was containing *43,000 records*, at the end of the project.

3. Diffusion

The database is currently *accessible on-line via DIMDI, DataStar and the computer host of Conseil Général of Alpes-Maritimes* (Videotex on Minitel). ICDB is monthly updated. *More than 600 users worldwide* are using the database. Due to the very high performance of host software, the connection time is very short and the net royalty income is low for the producers with this type of access.

Monthly updates of ICDB are also distributed on floppy disks.

4. Wider considerations

1. The Immunoclone Database is now the *biggest database on monoclonal antibodies and hybridomas* and the most accessible, worldwide. Procedures and organization can be easiliy adapted to similar projects.

2. This type of database related to the description of biological products may take advantage *in future to be closely linked with the delivery of the bioproducts* themselves. The user of the information is willing to get directly the product selected by the information search. It is easier, from the user point of view, to include the cost of information processing into the purchase of the product than to be asked for a separate payment of the information service. The traditional 'free' catalogue marketing technique may apply to scientific and technical information, as well.

3. This project pointed out *the need for a strategy in the field of biotechnology scientific information* (and probably in many other connected areas) at the Community level and at each national level. If such, it could improve its perenniality, make the cooperation with similar projects in Europe, America or Japan more efficient and less hazardous, increase the feed-back from the research community and improve its integration in research infrastructure.

MAJOR COOPERATIVE LINKS

1. Participation of all project members to the international conference BioInformatics in the '90s', thanks to DG XII/F organization (Maastricht, November 1991)

2. Three training sessions at Sophia-Antipolis in May 1991, September 1991 and January 1993, with attendees from each input centre.

3. The second International Workshop on Hybridoma and Animal Cell Lines, organized by CERDIC and ECACC(Sophia-Antipolis, April 27th, 1993 — 50 attendees).

4. Joined participation of ECACC and CERDIC to BioExpo 91 and 93 (Paris), and of ECACC,CERDIC and IST to Biotech 92 (Genova).

5. Dr. Alan Doyle, Curator of ECACC, was elected Chairman of the Board of CERDIC (December 1992-December 1993).

European dimension

Biological data collection and information transfer are typically international activities. *The European dimension is probably the minimum required for such projects.* There is no particular problem induced by a multinational partnership but in many circumstances it provides complementary experiences and know-how, e.g. the improvement of electronic communication, resource identifications, language translation, scientific analysis, etc.

1. Electronic communication and data access in research community

Many biomedical research institution equipment and know-how were below the minimum required for a reliable communication infrastructure. This was in progress during these last 2 years, thanks to the *success of Internet Protocol networks.*

One can expect that it will improve the transfer of scientific and technical knowledge in biology and increase the role of databases in this transfer. *It is important that those databases be validated and reliable sources of information, under scientific control.* The European scientific community may turn the integration of this activity in its research infrastructure to great advantage.

*2. **Biological databases in the European research community.***

The current budget of biological research laboratories cannot bear cost of scientific information produced on a commercial basis. Although such activity should be financed on a regular basis, most of the biological databases produced in Europe are in fact supported by special programmes, like BRIDGE. *The discontinuity of these programmes do not let any possibility to compete (or even cooperate on a equal basis) with foreign equivalent projects.* Some efforts are made for sequencing data but many other domains are concerned.

*3. **Availability of biotechnology products***

Some biological products, like hybridomas, cell line cultures, DNA probes have a strategic importance for the worldwide competition in pharmacology research, genetic disease investigations, etc. But the research bioproduct resources are much more scattered in Europe than in the US or in Japan. More and more bioproducts used by the European laboratories are of foreign origin because the ones which have been developped in Europe are unknown and consequently not internationnally recognized. *The support of information networks and database management in that domain may have a very positive effect on the development of biotechnologies in Europe.*

PUBLICATIONS

Monthly publication of *'Immunoclones'* (EC support aknowledged on the cover).

Electronic linking services for biotechnologists and microbiologists in Europe (BIOT CT-890042)

COORDINATOR:
E. ROSS, MSDN, Cambridge, UK
PARTICIPANTS:
MSDN is sole contractor.

BACKGROUND INFORMATION

A wealth of information relating to microbiology and related disciplines exists in Europe. This information is increasingly available in computerised form. In order to facilitate access to and use of such data the Microbial Strain Data Network (MSDN), operating from its Secretariat in Cambridge, UK, run a modern networking facility providing electronic mail, bulletin boards and databases, and providing a single access point to integrated microbiology, biotechnology and biodiversity data as well as a mechanism for encouraging communications with biologists worldwide. Such a network, supported by training services and help desk enables information on biological material to be easily obtained by European laboratories through a single menu system reducing the difficulties of locating data and microbial cultures necessary for biotechnology research and development.

The MSDN is a UK company limited by guarantee with an international committee of management. It is sponsored by four organisations of the International Council for Scientific Unions, the World Federation for Culture Collections (WFCC), the International Union of Microbiological Societies (IUMS), the Commission on Data for Science and Technology (CODATA), and the Scientific Committee for Biotechnology (COBIOTECH).

OBJECTIVES AND PRIMARY APPROACHES

MSDN's major aims in providing an integrated electronic information and communications network for microbiologists and biotechnologists are to: enhance access to existing data for scientists within the European Community by providing a single access point to relevant integrated data resources; encourage providers of specialist microbiology and biotechnology data to disseminate their information over the MSDN Network; support end-users of the MSDN Network through a help desk; facilitate communications between biologists both within Europe and beyond; and establish regional nodes of MSDN to provide local access, support and training.

MSDN Networked services are accessible through the public data networks (X.25) and the research-based Internet. The MSDN computer host is based at the Base de Dados Tropical (BDT), Campinas in Brazil. The switch to BDT as MSDN host was made from British Telecom in May 1993 in order to reduce costs in maintaining the MSDN Network, enable access through the Internet and allow the introduction of a simpler (cheaper for the end-user) method of charging. With the opening up of the Internet in recent years and the introduction of new Internetworking technologies it was important that MSDN databases were accessible through this route. Enhancements have been made to the MSDN Network's front-end allowing simpler and more effective searching by MSDN users. A cross-file searching facility allows a global search of all MSDN databases using a single command.

Electronic mail is used as the primary communications route for users of the MSDN Network. A directory of MSDN users is searchable. Database queries can be sent by electronic mail and this non-interactive approach is particularly useful for those without the means to make an interactive connection. Results of the search are returned to the user's mail box by email.

Databases are added in an ongoing fashion to the MSDN Network. The text retrieval software, INFO, developed and maintained at BDT is used to search MSDN databases. Databases cover microbiology, biotechnology and biodiversity. Information is provided on microbial strains, culture collections, biotechnology publications and meetings, and resources relating to the release of organisms into the environment. The scope includes cell lines, hybridomas, molecular probes and recombinant materials as well as microorganisms. Connections to other host computers on the Internet, such as the CRC's Human Genome Mapping Project (sequence data) and the Interlab databases of cell lines and molecular probes, allows use of these resources by MSDN users. Access to biodiversity databases on the BIN21 Network is provided through the MSDN Network. MSDN users can participate in online discussions on biodiversity (biodiv-l list) and issues concerning the releases of organisms into the environment (irro-l list).

MSDN is the Secretariat for the Information Resource on the Release of Organisms into the Environment (IRRO) and the network host for IRRO databases. A directory to worldwide information resources on environmental releases is being developed by the MSDN for the IRRO and will be publically available by August 1994.

A full list of MSDN databases is attached to this report.

MSDN provides training in an ongoing fashion either on an informal basis at the MSDN Secretariat or in a more formal teaching setting. MSDN also provides support to subscribers in their use of MSDN Network services. Information and instructional materials are kept up-to-date. The MSDN database instruction manual is available on disk as well as in printed form. The MSDN training manual used in MSDN workshops on the 'Use of computers in microbiologyhas been extensively revised.

The MSDN Network is being expanded to encompass regional Nodes who will provide networked services, support and training at a regional level. Agreements have been reached to date with the World Conservation Monitoring Centre (WCMC), Cambridge, UK and the Bioinformatics Distributed Information Centre (DIC), Pune, India to become MSDN Nodes. The WCMC will primarily provide services to European and North American users, the Bioinformatics DIC to users in Asia, and BDT to users in South and Central America. Other organisations are being identified to expand the MSDN Network into other regions of the world. For example, MSDN is negotiating with centres in Nairobi to provide an MSDN service for users in Africa. Similarly, discussions are being held with organisations in Moscow to provide services for MSDN users in Russia and countries of the former Soviet Union.

RESULTS AND DISCUSSION

MSDN networked services are used by public institutions, industry, non-governmental organisations, and individual researchers. Just under half MSDN's user base represent commercial organisations.

When MSDN's new host computer was introduced in May 1993 usage of networked services increased significantly. This reflects the improved service and also the fact that users are no longer charged an hourly connect fee as in the past. Users are now charged a fixed semi-annual subscription fee allocated depending on the category of user, i.e. whether individual, group or institutional. This means users can connect to the MSDN Network and download records with no additional charges except local telecommunications fees.

The change of host also meant that charges for storing data were no longer inhibitory (as with British Telecom). As a result, a significant increase in the number of MSDN databases was possible.

It is difficult to put a value on information, but judging from the use and interest in the data available through the MSDN Network, it is obvious that there is a need for data on microbiology and biotechnology and that given the ever growing mass of data available, it is vital that MSDN continues to make specialised data available through a single access point, raise public awareness about data available and provide training in its use. Through expansion of the MSDN Network by regional MSDN nodes, users are better served at local and regional levels, receiving support in their first language and making use of technologies most appropriate to local circumstances. African users of the MSDN Network, for example, will most likely use FIDONet as their communications route to MSDN. FIDO software together with high speed modems is especially suitable for regions of the world where telephone lines are poor or unreliable. Using automatic file compression, data or messages are sent and received at high speed to and from the host (or receiving) computer.

A priority continues to be training, and MSDN offers instruction in: use of MSDN services, development of databases, eg. to manage and analyse microbial data, and use of other biological resources, particularly those available over the Internet. By training information staff at centres throughout the world who are then responsible for educational activities, the numbers of people who are better informed and better able to understand microbiology and biotechnology data and networking will grow.

Networking tools and technologies have been improved and developed dramatically over recent years. With the emergence of, for example, gopher, World Wide Web, and Wais tools for locating Internet-based information, an incredible amount of information is relatively easily accessible through simple menus or pointer systems. MSDN will continue to make use of the latest networking tools and technologies wherever possible.

INDUSTRIAL APPLICATIONS

The MSDN Network makes available a unique specialized collection of data resources on microbiology and biotechnology. Pharmaceutical companies are a major MSDN user category. The different types of information available through the MSDN aid the biotechnologist or chemist in the development of a novel product or process. For example, databases can be used to identify a microbial culture with defined properties, determine which companies are working on biological control, assess biological diversity in a specific region to provide information on host-species interactions, etc.

MAJOR COOPERATIVE LINKS

As an international organisation, MSDN has forged links with many groups and individuals worldwide. BDT provide MSDN's database host computer and perform associated systems management functions. The WCMC is being established as the European MSDN node. The Bioinformatics DIC is being established as the Asian MSDN node. Within Eastern Europe, MSDN has an agreement to collaborate with the Institute of Physiology and Biochemistry of Microorganisms (IBPhM), Pushchino, Russia. As a result of this agreement, MSDN distributes many databases (principally microbial culture collection catalogues) originating from Russia and translated into English. MSDN also collaborates with the BIN21 Network by pointing to relevant biodiversity sources on BIN21 from the MSDN Network. Additionally, MSDN is the Secretariat for IRRO, and distributes and develops IRRO databases on releases of organisms into the environment.

PUBLICATIONS

Kirsop, B., Ross, E. and Nandi, S. A Biological Network: its conception, birth and evolution. Bioinformatics, Vol. 1 No. 3, 1992, pp. 26-30.
Kirsop, B., Ross, E. and Nandi, S. Development of a network for microbiologists: MSDN, a case history. Bioinformatics, Vol. 1 No. 2, 1992, pp. 3-10.

AREA B:

INFORMATION INFRASTRUCTURE

CULTURE COLLECTIONS

Information Centre for European Culture Collections (ICECC) (BIOT CT-900162)

COORDINATOR:

D. CLAUS, DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, D

PARTICIPANTS:

DSM is sole contractor

BACKGROUND INFORMATION

In 1981, the curators of the major service collections of biological materials in Europe were concerned that despite the existence of a number of important national culture collections within Europe, industry and the scientific community often were unaware of their existence or of the range of activities offered by these culture collections. As an important stage in the future development of European culture collections the coordination of the exchange of information between the collections and the provision of an information service to the users of the collections was proposed.

In recognition of these problems the Commission of the European Communities agreed in 1988 to establish an 'Information Centre for European Culture Collections' (ICECC) The project was started in the beginning of 1989 with full financial support from the Commission under its Biotechnology Action Programme. Subsequently, the ICECC project has been continued under the BRIDGE programme of the CEC with additional financial support coming from the host institution of the ICECC, the DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen in Braunschweig, Germany.

OBJECTIVES AND PRIMARY APPROACHES

The major tasks of the ICECC was to continue as the permanent secretariat for the European collections of biological cultures, to provide a central contact point for European scientists in all culture collection related matters and to act as a public relation and promotion office on behalf of the culture collections.

RESULTS AND DISCUSSION

The existence and the tasks of the Information Centre for European Culture Collections has been publicized from its beginning. As the promotion centre for European culture collections the ICECC has published since 1989 a series of Newsletters for which each of the European culture collection was asked to supply articles of interest and general collection news. These newsletters were distributed to European culture collections, to interested scientists, to companies and to administrations in the scientific area by post and at national and international scientific fairs and meetings in different European countries. The circulation number of these newsletters was as follows:

ICECC Newsletter
No. 1 (1989). Circulation: 1000 copies
No. 2 (1990). Circulation: 1500 copies
No. 3 (1991). Circulation: 2000 copies
No. 4 (1992). Circulation: 3000 copies
No. 5 (1993). Circulation: 2500 copies

At several international scientific meetings and fairs (Paris, Strasbourg, Hannover, Maastricht, Prague) the ICECC has shown its activities as well as the numerous services offered by European culture collections. Catalogues of the collections were presented and information on the deposit of patent strains was given. Many of the collections represented by the ICECC supplied their own information material for exhibition or distribution. For other collections this material was prepared by the ICECC itself in form of shorter leaflets. At all these meetings or fairs the ICECC was joined by at least one national collection of the respective European country, by a representative of the CEC supported Microbial Information Network Europe (MINE) and (partly) by the Microbial Strain Data Network (MSDN). For its presentations the ICECC has developed its own exhibition stand.

In 1992, the ICECC has published a 48 pages brochure under the title *European culture collections: Microbial diversity in safe hands*. This brochure (total number of copies: 3100 was mainly distributed in 1993 and includes information on all 43 member collections of the European Culture Collections' Organization. The holdings and services of each collection are described and a short introduction to the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedures and on the European facilities for the deposit of patent strains is given. Information on the tasks of the ICECC, on the European Culture Collections' Organization (ECCO), the Microbial Strain Data Network (MSDN) and the Microbial Information Network Europe (MINE) was also included in the brochure.

The ICECC publication *Instructions for Shipping Non-Infectious and Infectious Biological Substances* has found great interest and was published in 1993 in the 5th edition. On 36 pages the booklet provides detailed information on international regulations (post, airfreight) for the safe shipping of all kinds of biological material including genetically manipulated cells. Many practical packing aspects are included in the brochure. More than 250 copies have been distributed free of charge to industrial and scientific institutions, including all European culture collections, or to administrative offices.

In 1990 the ICECC started to publish its list *Bacterial Nomenclature Up-to-Date* in a printed form and, later, in a computerized version. The aim of the publication was to inform culture collections not only on newly described bacterial species but also to supply the latest valid name changes in bacterial systematics. The computerized version has been also found most useful in preparing the data on bacteria within the MINE project. The database was distributed quarterly free of charge to all European culture collections and to members of the UNESCO sponsored Microbiological Resource Centres (MIRCEN). To improve the use of bacterial names in scientific publications a free subscription has been offered also to the editors of the main microbiological journals in Europe. This offer was highly appreciated by most journals. To the scientific community the database is available for a fee.

Since its beginning a major task of the ICECC was to reply to enquiries from European and overseas biologists on the availability of specific strains of microorganisms (algae, bacteria, fungi, protozoa) or plasmids, of animal cell lines, plant cell cultures or plant viruses. During the years most enquiries received by the ICECC referred to bacteria (about 47% of total enquiries), followed by enquiries on fungi and yeasts (12%). In case the strain(s) were already deposited in a culture collection, the ICECC was usually in the position to locate the strain(s) in question at least in one European culture collection or in collections not located within Europe. This was possible after the ICECC has collected all culture collection

catalogues available within and outside Europe. Part of the catalogues were available in a computerized form, enabling ICECC quick strain searching. In all replies the ICECC — as a 'neutral institution' — has mentioned all collections maintaining the strain(s) requested. If an order was given to the ICECC the order was forwarded to the collection selected by the client himself.

The main other requests received by the ICECC were on microbial identification or chemotaxonomic methods offered by European culture collections (about 7%), on the preservation of cultures (5%), on patent deposits and related problems (6%), on microbial taxonomy (3%) or on microbial strains with specific degradation capabilities (2%) or were of general nature (18%). Enquiries received by the ICECC during 1993 are shown in the following table:

Enquiries from industrial and scientific institutions in 1993 (except Shipping List, Bacterial Nomenclature-Up-To-Date, Price List, ECCO brochure, Newsletters)

Total:	481
Industry:	255
Science:	226

No. of enquiries from different countries

Europe

Austria	11	Russia	12
Belgium	13	Slowenia	3
Croatia	2	Slovakia	2
Czech Republic	9	Spain	14
Denmark	8	Sweden	11
Estonia	2	Switzerland	22
Finland	9	Turkey	9
France	36	United Kingdom	64
Germany	85	Ukraine	2
Greece	9		
Island	2	*Other countries*	
Italy	22		
Lithuania	2	North America	34
Netherlands	28	South America	12
Norway	5	Africa	6
Poland	10	Asia	22
Portugal	6	Australia	9

As in the years before the ICECC has published also in 1993 a new price list of the European culture collections.

During 1993 the ICECC has finished an investigation on the deposit of microorganisms used in scientific publications in public culture collections. This study, which will be published elsewhere in more detail, has shown the following results:

Microbiological journals studied (only 1992 volumes)	17
Number of publications studied	1985
Total number of strains used in these publications	8304
Number of strains received from a service collection	2159 (=26%)
Number of strains not received from a service collection	6145 (=74%)
Number of strains deposited in a service collection (without strains received from a collection)	**581 (= 7%)**

According to the survey 26% of the microbial strains used in scientific studies are coming from service collections. This figure confirms the important role service culture collections play in scientific research. However, only 7% of the other strains are deposited in service collections. This means that about 65% of the microbial strains of all scientific studies may not be available to other scientist or may be lost in future. This problem must be discusssed by the service culture collections and their organizations but is of general importance also for all funding organizations.

MAJOR COOPERATIVE LINKS

The Information Centre has established good connections to all European service culture collections and to the main collections outside Europe. Since 1992 new contacts have been sought to culture collections from middle and eastern European countries. Some of these collections were rather unknown in western Europe but may considerably contribute to the microbial diversity held in European culture collections.

Since 1990 the ICECC offered the database of the UK culture collections, the Microbial Culture Information System (MiCIS), after the UK government has stopped its funding. Searching of MiCIS information was mainly done via the Microbial Strain Data Network. After most of the MiCIS data has now been transferred to the MINE database, it was decided by the DSM not to continue to offer the MiCIS database after the termination of the ICECC. The ICECC itself was strongly engaged in the transfer of bacterial data of MiCIS and of other culture collections to the MINE database.

The secretary of the Information Centre has been elected by the members of the European Culture Collections' Organization (ECCO) as the secretary of this organization. In this function she was engaged in the preparation of the annual ECCO meetings in Valencia (1991), Göteborg (1992) and Istanbul (1993).

Within the scope of the annually held International Training Programme, given by the GBF in Braunschweig, the ICECC presented to the participants lectures on its tasks and on the Microbial Information Network Europe (MINE).

In 1993, the ICECC was involved in the organization of two workshops on *Advanced Techniques in Animal Cell Cultures and Techniques in Plant Cell Cultures*. Both workshops were held during May in Brno, Czech Republic, by staff members of the Department of Biology, University of Brno, the European Collection of Animal Cell Cultures, and the DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen. Participants were coming from countries from eastern and western Europe. The workshop on Animal Cells was supported by the Commission of the European Communities. Besides the organization of the workshops and public relations activities, the course manuals for both workshops were prepared by the ICECC.

Conclusions

In the beginning of the project, the ICECC had a lot of difficulties in developing its tasks within the CEC culture collections due to certain objections coming from certain member collections of the Microbial Identification Network Europe (MINE) but not from other CEC member collections of the European Culture Collection's Organization. The former collections considered the ICECC mainly as a competitor for funds available for European culture collections under the BAP/ BRIDGE programme of the CEC. The situation changed when the ICECC started

to present annual reports on its activities during the annual meeting of ECCO (first report was given 1990 during ECCO IX in Valencia). During the XIIth ECCO Meeting held in Istanbul in June 1993 the ICECC was considered as an important institution in the further development of European culture collections. Proposals to continue the ICECC after the end of the support by the CEC and the DSM have been discussed. ECCO was not able to come to a final decision. In the meantime it was agreed that the project leader of the present ICECC will continue to run the Information Centre at an reduced level. The DSM has agreed to continue to house the ICECC in case ECCO is able to find financial support for the further running of the centre.

PUBLICATIONS

ICECC publications (1993): Instructions for Shipping Non-Infections and Infections Biological Substances.
ICECC publications (1992): European culture collections: Microbial diversity safe hands.

Microbial Information Network Europe (MINE) (BIOT CT-910280)

COORDINATOR:
D. van der MEI, CBS, Baarn, NL
PARTICIPANTS:
D.L. HAWKSWORTH, Egham, GB
J. de BRABANDERE, Brussels,B
G. KALANTZOPOULOS, Athens, GR
I. SEDLÁCEK, Brno, CZ
A. MARTINI, Perugia, I
F. URUBURU, Burjasot,E
I. SPENCER-MARTINS, Oeiras, P
Y. MENORET, Paris, F
E. FALSEN, Gothenburg, S
E. STACKEBRANDT, Braunschweig, D
M.-L. SUIHKO, Aspoo, SF

BACKGROUND INFORMATION

It was acknowledged by the EC that the diverse and rich genetic resources present in collections of microorganisms in Europe should be made more easily accessible to the increasing population of potential users in biodiversity research, biotechnology, agriculture, management and conservation of nature, healthcare, waste management and the use of natural resources for a sustainable future.
During the Biotechnology Action Programme (BAP), a common format for storage and retrieval of data was devised, as a first and major step to harmonize the presentation of the data.

The results were presented in two papers:
W. Gams et al (1988): Structuring strain data for storage and retrieval of information on Fungi and Yeasts in MINE, the Microbial Information Network Europe: Journal of General Microbiology, 134, 1667-1689
and:
J.A. Stalpers et al (1990): Structuring strain data for storage and retrieval of information on Bacteria in MINE, the Microbial Information Network Europe: Systematic and Applied Microbiology, 13, 92-103.

The criteria for adoption of (new) Member Collections in the MINE-system were presented in:
D.L. Hawksworth and M.A.A. Schipper (1989): Criteria for consideration in the accreditation of culture collections participating in MINE, the Microbial Information Network Europe: Mircen Journal, 5, 277-281.

The format for the MINE-programme, as used during the integration, consists of the Minimum Data Set (MDS), which contains the general information and the Full Data Set (FDS) in which more specialized and detailed data are included.

Fields used for MDS and FDS.

Names of species and authors
MDS:

SP	Species name and authors
SSP	Subspecies name and authors
VAR	Variety name and authors

PVAR	Pathovar and authors
F	Form name and authors
FSP	Special form name and authors
RACE	Race name and authors
MOR	Morph (anamorph or teleomorph)
MIS	Misapplied names and authors
OTHNAM	Other names and authors

FDS:

TAX	Taxonomy (names of suprageneric taxa)
SERO	Serotype, serovar
TAXREM	Taxonomic remarks
LIT	Taxonomic literature
SUPPLY	Form of supply to the collections

Strain administration

MDS:

STN	Strain number
OCC	Numbers in other collections
EDA	Date of accession
RESTR	Restrictions and precautions
STAT	Status (indicating type or authentic strains)

FDS:

REM	General remarks
PRE	Mode of preservation

Environment and history

MDS:

ISOFR	Source from which the strain was isolated
LOC	Location of original material
ISOL	Name of isolator of the strain
DET	Name of microbiologist who identified the strain
DEP	Depositor of the strain in the present collection
HIS	History of the strain between isolation and deposit

FDS:

SUSP	Substratum specifity for species
HAB	Habitat (biotope)
COLL	Collector of original material
HERB	Herbarium where specimen is deposited
ISOM	Isolation method
ENHISLIT	Literature on environment and history
ENHISREM	Remarks on environment and history

Sexuality

MDS

SEX	Sexual behaviour of the species
SEXST	Sexual state of the strain

Biological interactions

MDS:

PATH	Pathogenicity of strain or species

FDS:

SYMB	Symbiotic relations of the strain or species
MPAR	Mycoparasitic activities of the strain or species

TOX	Toxicity to other organisms
ANT	Antagonistic activities against other organisms
BINTLIT	Literature on biological interactions
BINTREM	Remarks on biological interactions

Properties (cytology, biomolecular data)

FDS:

NUCL	Number of nuclei in species or strain
WCONS	Wall constituents
CHTAX	Chemotaxonomical markers, including co-enzyme Q
GRAM	Gram behaviour
STAIN	Staining reactions
PIGM	Pigments
GC	Guanine-cytosine contents of DNA
TKIT	Result of test kits
PROPLIT	Literature on properties
PROPREM	Remarks and properties

Genetic data

MDS:

MUT	Type of mutant
KIL	Killer properties of yeasts

FDS:

MUTMET	Mutation method
GENOT	Genotype (chromosomal markers)
HYBR	Hybrids
PLS	Plasmids occurring in the strain
PHAGE	Phages occurring in the strain
GENLIT	Literature on genetic data
GENREM	Remarks on genetic data

Growth conditions

MDS:

CONDS	Conditions for growth on solid or liquid media
TEMPR	Temperature relationships of species (psychophilic, etc.)

FDS:

CONDSP	Conditions for fruiting and sporulation
CONDGER	Conditions for germination
NUGR	Nutritional requirements and growth factors
TOL	Tolerances and sensitivities
DEF	Deficiencies
METY	Metabolic type
CNUT	Carbon nutrition type
CSOR	Single carbon sources tested for growth
NSOR	Single nitrogen sources tested for growth
CNSOR	Single compounds tested as sole C and N sources for growth
OXR	Oxygen relationships (aerobic, etc.)
TEGR	Cardinal temperatures for growth
TESP	Optimum temperature for sporulation
HEATR	Heat resistance
PHR	pH relationships (acidophilic, etc.)
PHC	pH conditions for growth
SALR	Salinity relationships (halophilic, etc.)
SALC	Salinity conditions for growth

SUGC	Optimum and maximum sugar concentrations
ELDO	Electron donor
ENERS	Energy source
ELAC	Special electron acceptors
CONDLIT	Literature on conditions
CONDREM	Remarks on conditions

Chemistry and enzymes
FDS:

ENZ	Enzymes produced
MET	Metabolites produced
DEC	Decomposition, degrading and utilization capacities
LEACH	Leaching
BIOLUM	Bioluminescence
CHLIT	Literature on chemical data
CHREM	Remarks on chemical data

Practical applications
MDS:

APPL	Industrial and general applications
TEST	Applications in testing and assays
PATENT	Information on patents

OBJECTIVES AND PRIMARY APPROACHES

The objectives of the MINE-project were twofold:

- To establish a common format for registration of data on strains of microorganisms (bacteria, fungi and yeasts) present in European collections.
- To facilitate access to these data, mainly through the foundation of a Central Data Node.

The project involves 40 collections in 12 European countries; Sweden, Finland and Republic Czech participated without support of EC.

The coordination at the national level was achieved through National Nodes, often one of the main collections, sometimes another representative body. The representatives of the National Nodes constituted, as project managers, the Board of the project, which was presided by the coordinator.

The contract was concluded between the EC and the Royal Netherlands Academy of Arts and Sciences (RNAAS) as main contractor, all other parties being associated contractors.

Harmonization of data as well as taxonomic and nomenclatural issues, were dealt with through Responsible Committees (RC's), one for bacteria (chaired by D. Janssens, Ghent, Belgium), one for fungi and one for yeasts (both chaired by G. Hennebert, Louvain-la-Neuve, Belgium).

Data Integrating Nodes (DIN's) were assigned to the task of integration and coordination of data input by Member Collections, and the transfer of data to the Central Data Node (CDN). The DIN for bacteria was established at the Laboratorium voor Microbiologie en Microbiële Genetica (LMG); that for fungi and yeasts at Centraalbureau voor Schimmelcultures (CBS).

The CDN was established at the Deutsches Institut für Medizinische Dokumentation und Information (DIMDI).

RESULTS AND DISCUSSION

A. Implementation

The implementation of the MINE-programme, culminating in the realization of the Central Data Node, took a long preparative phase. Two items proved to be particularly cumbersome: the form and choice of the CDN and the financing of the project.

Only in the end of 1991 the definite choice was made to host the database at DIMDI.

The financial paragraph involved lengthy discussion with EC about the support for the project and — predictably, considering the large amount of collections and countries — the division of this support among the participants as Member Collection, Data Integrating Node, in coordination and as members of Responsible Committees. The key for retribution was established during a meeting of project managers in Paris, 1989.

The need for an information centre to inform users about the potential value of collections and the CDN, was acknowledged by EC. For this a separate project was concluded between EC and DSM. The transition of BAP to MINE was a gradual one, work on the implementation of MINE already starting during the interlude of nearly two years between the end of the BAP programme and the onset of the MINE programme. Member Collections developed their databases for entering data according to the common format, and a first integration of the MDS was performed by the DIN's.

At the meeting of project managers in Valencia the official starting date for the project was fixed at June 1, 1991.

B. MDS

From June 1, 1991 on the DIN's went full ahead with the integration of the MDS. Towards the end of 1991 this integration was completed for bacteria as well as for fungi and yeasts. For correcting and harmonization, reports of the integrated MDS were prepared for the RC's.
Correcting and harmonization of the MDS for Fungi and Yeasts was finished by the RC's towards the end of 1992; the data on unique strains were entered in the CDN in February 1993, those on duplicate strains in June 1993.

Correcting and harmonization for Bacteria took more time; at the moment 90% of all strain data have been entered at the CDN, the last 10% is expected to be entered towards November 1994. The CDN went into operation in July 1993.

For correct naming of species, the approved list of bacterial names was used; a computerized version of this list is maintained and made available by DSM.

For fungi and yeasts such a list was not yet available. A reference list was created by CBS with full author names and synonym- and anamorph/teleomorph relations. The list now contains 19.215 species names.

C. FDS

In view of the delay in the realization of the MDS, work on the FDS was started simultaneously. In 1993, the discussion about fields to be covered and harmonization of the FDS was intitiated. For fungi and yeasts the Member Collections concerned sent their data between March and June 1994 to the DIN, where they were subsequently loaded and integrated. They are now up to correction by the RC's, and it is envisaged to enter them at the CDN towards November 1994.

Work on the FDS for bacteria will start after finishing the MDS.
The experiences gained during the first integration will greatly facilitate the procedure for correction and harmonization, the main factors being:

— decisions taken by RC's are saved and re-used;
— strict adherance to standards, which diminishes the number of necessary corrections;
— corrections are performed automatically by the DIN's;
— the 'field by field' approach used in the FDS is more efficient than the originally applied 'record by record' approach.

After receiving the corrected data for fungi and yeasts, the CDN will start work on the menu for the FDS.

NUMBER OF STRAINS IN THE MDS		*Fungi/Yeast*	Bacteria
TOTAL NUMBER		58.538	32.749
PRESENT IN ONLY ONE COLLECTION (UNIQUE STRAINS)		45.159	22.111
PRESENT IN AT LEAST TWO COLLECTIONS (DUPLICATE STRAINS)		13.379	10.638
CONTRIBUTION OF MEMBER COLLECTIONS			
B	IHEM	1.896	
	MUCL	6.108	
	LMG		7.211
FR	LCP	1.130	
	ANIE		21
	CFBP		1.455
	CNRZ		617
SF	VTT	566	
FRG	DSM	1.817	3.800
GR	ATHUM	216	
	BPIC	332	1.066
	ACA-DC		78
I	DBVPG	2.115	
NL	CBS	27.454	937
	KIT		209
	LMAU		695
	LMD		1.024
	PC		2.790
	PD		1.155
P	IGC	451	
E	CECT	874	1.113
S	UPSC	1.740	
	CCUG		1.518
UK	IMI	10.576	
	NCPF	823	
	NCYC	1.886	
	NWRF	554	
	NCFB		2.119
	NCIB		3.462
	NCMB		816
	NCTC		2.663

Number of strains in the FDS		*Fungi/Yeasts*
Total number		*64.810*
B	IHEM	1.895
	MUCL	8.917
FR	LCP	1.127
	CLIB	152
SF	VTT	586
FRG	DSM	1.947
GR	ATHUM	312
	BPIC	341
I	DBVPG	2.303
NL	CBS	28.707
P	IGC	919
E	CECT	847
S	UPSC	2.044
UK	IMI	11.470
	NCPF	820
	NCYC	1.869
	NWRF	554

MAJOR SCIENTIFIC BREAKTHROUGHS

The MINE project achieved a firm cooperation between 40 European Collections, with as a major breakthrough the establishment of a common format for handling of data of microorganisms.

Although the CDN is not yet fully operative, it will be at the end of 1994 for the Fungi and for the MDS of Bacteria; the FDS for Bacteria is expected to become available early 1995.

The number of catalogues produced during the MINE project, along the lines outlined in the MDS, amounted to 25; many catalogues were first printings, others updates form rather antiquated ones. This means a major contribution to the dissemination of information in collections, the catalogue, apart from on-line systems, still being an important vehicle in this respect.

Moreover, at the national level, collections were given a more firm constitution, for example, in Belgium and the Netherlands.

It was unanimously recommended by the project managers, to continue the cooperative effort with emphasis on data-handling, formatting, harmonization and exploitation of the CND.
With support of the VALUE Programme the outline for further cooperation in the form of a consortium was presented during the last project managers meeting at Louvain-la-Neuve, July 1994. It is the general feeling that this outline, with certain adaptations, can provide a good framework. In the coming months the statues will be further commented upon.
It is clear, although EC has not yet made any future commitments to the MINE effort, support for special projects, aiming at the further perfection and completing of the system, also in view of the fast developments in the fields of networking (INTERNET) is essential for the succezss of this endeavour.

MAJOR COOPERATIVE LINKS

1991	Valencia, E	Project managers meeting; start of the project.
1992	Brussels, B	Meeting of RC's, DIN's and CDN.
1992	Cologne, D	Meeting of MC's, DIN's and CDN.
1992	Gothenburg, S	Project managers meeting.
1993	Egham, GB	Project managers meeting.
1993	Baarn, NL	Meeting of RC's, DIN's and CDN.
1994	Brussels, B	MINE after BRIDGE.
1994	Louvain-la-Neuve, B	Project managers meeting

Cooperative links were established between the European Culture Collections Organization (ECCO) and the World Federation for Culture Collections (WFCC).

PUBLICATIONS

Joint publications

Gams, W., Henneber, G.L., Stalpers, J.A., Janssens, D., Schipper, M.A.A., Smith, J., Yarrow, D. and Hawksworth, D.L. 1988. Structuring strain data for storage and retrieval of information on fungi and yeasts in MINE, the Microbial Information Network Europe. J. Gen. Microbial. **134**: 1667-1689.

Hawskworth, D.L. and Schipper, M.A.A. 1989. Criteria for consideration in the accreditation of culture collections participating in MINE, the Microbial Information Network Europe. Micren Journal. **5**: 277-281.

Stalpers, J.A., Kracht, M., Janssens, D., de Ley, J., van der Toorn, J., Smith, J., Claus, D. and Hippe, H. 1990. Structuring strain data for storage and retrieval of information on bacteria in MINE, the Microbial Information Network Europe. Syst. Appl. Microbiol. **13**: 92-103.

Individual publications

Gams, W., Stalpers, J.A. Stegehuis, G.J. and Smith, J. 1990. Computerization of strain data in the Microbial Information Network Europe (MINE). Sydowia **42**: 218-230.

AREA B:

ENABLING TECHNOLOGIES

PROTEIN DESIGN/MOLECULAR MODELLING

Engineering of microbial peptide lantibiotics for use in agro-food and biomedical industry (MNLA) (BIOT CT-910265)

COORDINATOR:
(A) C.W. HILBERS, KUN, Nijmegen, NL
PARTICIPANTS:
(A) H.-G. SAHL, Fr. Wilhelms Univ., Bonn, D
(B) R. SIEZEN, Ned. Inst. Zuivel Onderzoek, Ede, NL
(C) G. JUNG, Eberhardt-Karls-Univ., Tübingen, D

BACKGROUND INFORMATION

Lanbiotics are a familly of antimicrobial peptides that contain the rare amino acid lanthionine. They are produced by Gram-positive bacteria. The most prominent one, nisin, from *Lactococcus lactis*, is employed as a preservative in food industry. Epidermin and mersacidin are potential therapeutic agents against bacterial diseases, the lantibiotics of the duramycin group are inhibitors of eukaryotic enzymes.

OBJECTIVES AND PRIMARY APPROACHES

The project was aimed at the elucidation of the structure-function relationships of lantibiotics and at the construction of novel lantibiotics with improved properties concerning stability, solubility, target organism, etc. To this end the following objectives were set:

(i) isolation and testing of novel lantibiotics from natural sources,
(ii) construction and testing of mutant lantibiotics,
(iii) elucidation of the three-dimensional structures of lantibiotics, and
(iv) elucidation of the processes involved in lantibiotic biosynthesis, posttranslational modification and secretion, and of the process involved in cell killing and self-immunity.

RESULTS AND DISCUSSION

A. Novel lantibiotics from natural sources

A total of 800 staphylococcal strains were screened for production of lantibiotics (B). Several new lantibiotics have been purified and their structure was elucidated (D) (Pep280, several epidermin analoga, SA-FF22 and carnocin U149). Epilancin K7 was purified on a large scale (B) for elucidation of the primary and three dimensional structure (A). Besides it could be shown that staphylcoccin 1580 is epidermin. The screened nisin-producing *L. lactis* strains were found to produce either the variant nisin A or nisin Z, which differ in position 27 (His → Asn). A second lantibiotic from *L. lactis*, lactacin 481, was identified (C) and characterized (collaboration with J.C. Picard, INRA, Jouy-en-Josa, F).

B. Lantibiotic biosynthesis

About 10 kb of the plasmid pED503 that codes for Pep5 production in the producer strain *S. epidermis* 5 have been sequenced and revealed the Pep5 biosynthesis cluster, which comprises a transporter gene *(pepT)*, an immunity gene *(pepI)*, the structural gene *(pepA)*, a putative protease *(pepP)*, and two biosynthesis genes *(pebB* and *pepC)*. Finally, it could be shown that the transporter is not necessary for Pep5 production, whereas the protease gene has to be present. Using antibodies raised against a fusion protein of *pepI* linked to the maltose binding

protein it could be shown that all phenotypically immune clones express the PepI protein. A main goal was the structure elucidation of intermediates in the biosynthesis of lantibiotices such as Pep5 (B), which were found to arise from a precursor protein consisting of a leader sequence followed by the Pep5 part in different stages of dehydrated serine and threonine residues as became clear from sequence analysis and electrospray mass spectrometry (D). All biosynthetic intermediates were found in the cytoplasmic cell fraction (B). This implics that all steps of the Pep5 biosynthesis take place in the cytoplasm.

Analysis (C) of the nisin operon on transposon Tn5276 has shown that seven genes are located directly downstream of the structural nisin A gene. These genes (and encoded products) are named *nisT* (translocater protein), *nisB* and *nisC* (possibly modifying enzymes, *nisI* (immunity protein), *nisP* (leader peptidase), *nisK* and *nisR* (regulatory proteins). All genes have been sequenced completely. Gene disruptions in *nisB, nisC* or *nisIPA* lead to reduced immunity and absence of nising production. Disruption of *nisIPA* with complementation of *nisA* leads to the production of precursor nisin A.

The gene for the novel lantibiotic epilancin K7 was localized on the chromosome of *S. epidermis K7* by Southern hybridization (A, B). DNA probes were designed on the basis of the amino acid sequence, derived by NMR (A). A gene-harbouring DNA fragment was cloned and sequenced (A, C). In addition to the structural gene *(elkA)* partial ORFs were found, crossing the borders of the fragment, named *elkP* and *elkT*, coding for a putative protease and translocator. The *elkA* sequence corroborated the NMR-derived sequence.

By comparison of the sequences of the genes involved in nisin, Pep5 and epilancin K7 biosynthesis, also including available sequences (subtilin, epidermin) from the literature, a clear picture of the proteins involved in biosynthesis and self-immunity of type-A lantibiotics is emerging. By computer-simulation (C) the leader peptidases NisP and EpiP were modelled.

C. Mutant lantibiotics

Site-directed mutagenesis of PepI (B) has demonstrated that the immunity is decreased when the hydrophobic sequence in the N-terminal part of the peptide is interrupted. An expression system for site-directed mutagenesis was constructed in the vector pCU1 which allowed the production of modified Pep5 molecules in a variant of the wildtype producer strain that had been cured from the plasmid coding for bacteriocin production (B). The modified peptides were designed on the basis of the spatial structure (D) of Pep5 in a membrane-mimicking solvent; each was purified to homogeneity (B) and modifications confirmed by mass spectrometry (D). It became clear that the modified residues (lanthionine rings, dehydrated amino acids) as well as the amphiphilic properties and the central flexible region are indispensable for activity. Several different systems have been developed (C) for the expression of mutant nisin genes, one of which also produces nisin A in addition to mutant nisin. By site-directed mutagenesis several tens of nisin A and nisin Z mutants were generated (C), and characterized by NMR (A, C), activity and solubility assays (C). E.g. the solubility of nisin at neutral pH was improved by sustituting His27 or His31 for lysine. Several mutations at position 5 were performed, since Dha5 plays an essential role in activity of nisin; other mutations were in the important flexible linker-region connecting lanthionine ring C with D/E. More than ten mutations in the leader sequence of nisin have been made (C) to study the role of the leader in biosynthesis, secretion and processing. Several precursors of nisin have been produced and characterized.

A phenotypic (activity) screening assay was developed (C) for randomly mutagenized *nis* genes expressed in *L. lactis* NZ9700 and NZ9800, one of the target organisms is *Listeria monocytogenes.*

D. Chemical Synthesis

All chemical syntheses were carried out in (D). The synthesis of the leader peptide (23aa), pro-nisin (34aa) and pre-nisin (all 57aa) was performed. The synthesis of such a long peptide, containing five cysteines, is still a major achievement, even for experienced peptide chemists. It was successful via two approaches: the condensation of three fragments and the stepwise synthesis in full length. Synthetic peptides were also produced for antibody production and detection of the *nisA* gene product (C) and Pep5 precursor protein (B).

A novel route to the solid state synthesis of peptides containing the unsaturated residues Dha and Dhb has been elaborated. The total synthesis of the immunity peptide of Pep5 (71aa) has been carried out in amounts and purity sufficient for biological studies of its potential activity. The polypeptide has a calculated molecular mass of 8070, which has been experimentally verified via electrospray-MS. It has unique amphiphilic properties, according to CD the N- and C-terminal segments adopt α-helical structures and the middle part consists of β-sheet conformation (D).

A chromophore-labeled substrate for the leader peptidase of nisin was prepared, used (C) for screening for this enzyme.

E. Structures of lantibiotics

Many mass spectrometric analysis and protein sequencing studies were carried out (D). These included isolated precursors of Pep5 (B) and various nisin precursors including a nizin Z with a subtilin leader sequence (C). Studies on degradation products of nisin (C) revealed important information on the critical residues of this bacteriocin.

Novel derivatization procedures, prior to Edmann degradation and Mass spectrometry, were developed (D) which allow for the first time an access to:

(a) N-termially blocked lantibiotics,
(b) full sequencing via Edmann degradation,
(c) mass spectrometric determination of unsaturated α,β-didehydro residues and thioether bridges,
(d) oxidation of thioethers and reduction of sulfoxides in lantibiotics.

The conformations of natural pre-Pep5, synthetic prenisin and of all leader peptides of prelantibiotics were investigated by circular dichroism in lipophilic solvents (D). All exhibited essentially α-helix structure, which was most likely induced by the lipophilic environment; the nisin leader and pre-nisin appear random coiled in aqueous solution as indicated by NMR (A).

Electronspray MS of prenisin revealed that it formed a complex of prenisin with two Zn^{2+} ions. This specific metal binding may be a prerequisite for correct folding prior to enzymatic modification and may prevent cysteine thiol groups from oxidation.

Both the primary and spatial structure of the novel lantibiotic epilancin K7 were unravelled *via* NMR (A). This is a novel approach; normally NMR studies require the amino acid sequence to be known *a priori*. Three-dimensional ^1H NMR and

heteronuclear (^1H,^{13}C) NMR were found very useful to solve the structure. The results showed that K7 is a novel lantibiotic indeed: it contains two Dha's, two Dhb's and two MeLan's, which is a composition different from any other known lantibiotic. The correct pairing of the MeLan thioether bridges could be established.

The three dimensional structures, in aqueous solution, of nisin (A), gallidermin (D) and Pep5 (D) were determined by NMR. Although the structures appear rather flexible, some common characteristics could be discerned. They posses a screw-like shape, with a clearly ampiphilic character, i.e. along the long axis of the molecule there is one typically hydrophobic face and one less hydrophobic. These findings are in agreement with a model in which lantibiotic molecules penetrate the lipid bilayer.

The structure of lacticin 481 was studied by 2D NMR (Collaboration with Dr. J.C. Picard, INRA, F) (C, A). NMR studies have been conducted (A) to unravel the structure of nisin in a membrane-like environment. It was found that nisin complexes to micelles of the detergents DPC and SDS and assumes a structure similar to the one in aqueous solution. The most prominent difference is a conformational change involving the residue, essential for activity, Dha5. The relative orientations of nisin and micelles have been characterized using micelles containing electron spin probes.

MAJOR SCIENTIFIC BREAKTHROUGHS

- Isolation of novel lantibiotics from natural sources.
- Production of site-directed mutants of Pep5 and nisin.
- Complete sequence of the Pep5 and nisin gene-clusters and functional analysis of the individual genes.
- Identification of the Pep5 immunity mechanism.
- New methods for aa-sequencing of lantibiotics.
- Spatial structures of several lantibiotics resolved.

MAJOR COOPERATIVE LINKS

Epilancin K7: isolated in B, NMR structure determination in A, isolation and sequencing of gene-cluster in B and C. This was the project of one scientist, stationed in A who spent one month in B and two months in C.

Nisin mutants/variants: purification in C, NMR in A and D. To this end one scientist, stationed in C spent a few days per month in A.

Pep5 mutants/variants and other novel lantibiotics: isolated in B, purification and chemical structure elucidation in D. This involved short visits of coworkers from B to D.

Synthetic peptides for immunizations: prepared in D, used in B and C.

Vectors, sequence data, molecular-biology methods: exchange between B and D.

Several meetings of all scientists involved took place in Bonn, Ede, Nijmegen and Tübingen.

The first international workshop on lantibiotics (Bad Honef, D, 1991) was organized by B and D; the second one (Papendal, NL, 1994) is organized by A in cooperation with B, C and D.

PUBLICATIONS

Joint publications

In: G. Jung and H.-G. Sahl (eds.), *Nisin and novel lantibiotics*, ESCOM, Leiden, 1991. Freund et al., p. 103-112; Kellner et al., p. 141-158; Frey et al. p. 180-188; Sahl et al. p. 332-346; Benz et al. p. 359-372; De Vos et al. p. 457-463.

T. Kupke, S. Stevanovic, H.-G. Sahl and F. Götz (1992). J. Bacteriol. **174**, 5354-5361.

H.W. van den Hooven, F. Fogolari, H.S. Rollema, R.N.H. Konings, C.W. Hilbers and F.J.M. van de Ven (1993). FEBS Lett. **319**, 189-194.

R. Jack, A. Carne, J. Metzger, S. Freund, H.-G. Sahl, G. Jung, J. Tagg (1994). Eur. J. Biochem, (in press).

M. v.d. Kamp, H.-G. Sahl, F. v.d. Ven. *Structure of epilancin K7* (in preparation)

M. v.d. Kamp, H.-G. Sahl, F. v.d. Ven, O. Kuipers. *DNA sequence epilancin K7 gene* (in preparation).

S. Freun d, H.-G. Sahl, G. Jung. *Structure of Pep 5 in solution* (in preparation).

H.-G. Sahl, R. Benz, G. Jung. *Membrane depolarization by epidermin and gallidermin* (Biochim. Biophys. Acta, to be submitted).

G. Bierbaum, J. Metzger, J.G. Jung, H.-G. Sahl. *Structure-function relationship in Pep5* (in preparation).

H.-G. Sahl, I.G. Bierbaum, J.J. Metzger, S. Stevanovic, G. Jung. *Isolation of natural epidermin variants* (in preparation).

M. Pag, M. Reis, G. Bierbaum, S. Surovoy, G. Jung, H.-G. Sahl. *Functional analysis of PepI coverring immunity to the producer of lantibotic Pep5* (in preparation)

Individual publications

In: G. Jung & H.-G. Sahl (eds.), *Nisin and novel lantibiotics*, ESCOM, Leiden (1991). F.J.M. van de Ven et al., pp. 35-42; M. Reis & H.-G. Sahl, pp. 320-331; H.-G. Sahl, pp. 347-358; G. Bierbaum & H.-G. Sahl, pp. 386-396; E. Molitor & H.-G. Sahl, pp. 434-439; G. Jung, pp. 1-34; S. Freund et al., pp. 91-102; A. Beck-Sickinger et al., pp. 218-230; O.P. Kuipers et al. pp. 250-259; P.J.G. Rauch et al., pp. 243-249; H.S. Rollema et al., pp. 123-130.

In: C. Schneider & A. Eberle (eds.), *Peptides 1992*, Proc. 22nd Pept. Symp., Escom, Leiden (1993). G. Videnov, S. Stevanović, T. Kupke, F. Götz, H.-B. sahl, pp. 938-939; H.A. Remmer, G. Jung, P. Rücknagel & E. Jaeger, pp. 167-168; A. Surovoy, D. Waidelich, G. Jung, pp. 563-564; S. Freund, N. Zimmermann, A. Fredenhagen, G. Jung, pp. 519-520.

In: R. James, C. Lazdunski, F. Plattus (eds.), *Bacterial plasmid-coded toxins: bacteriocins, micrococci and lantibiotics*, NATO ASI Series H: Cell Biology, Vol. 65 (1992). H.-G. Sahl: *Lantibiotics, an introduction*; H.-G. Sahl: *Biosynthesis of the lantibiotic Pep5 and mode of action of type A lantibiotics*; R.N.H. Konings et al.: *NMR studies of lantibiotics. The 3D structure of nisin*; S. Freund & G. Jung, pp. 75-92.

G. Jung, Angew. Chem. Int. ed. Engl. **30**, 1051-1193 (1991).

J. Hugenholtz, O.P. Kuipers, H.S. Rollema. Voedingsmiddelentechnologie, **21**, 19-22, (1991]

O.P. Kuipers; H.J. Boot, W.M. de Vos, Nucl Ac. Res. **19**, 4558 (1991).

J.W.M. Mulders, I.J. Boerrigter, H.S. Rollema, H.S. Siezen, R.J. de Vos, Eur. J. Biochem., **201**, 581-584 (1991).

P.J.G. Rauch, W.M. de Vos, J. Bacteriology **174**, 1280-1287 (1992).

F.J.M. van de Ven, H.W. van den Hooven, R.N. Konings, C.W. Hilbers, Eur. J. Biochem. **292**, 1181-1188 (1991).

G. Jung & A.G. Beck-Sickinger (1992): Angew. Chem. **104**, 375-391; Ang. Chem. Int. **31**, 367-383.

A.G. Beck-Sickinger, N. Zimmermann, G. Jung (1992) in: R. Epton (eds.), *Innovation and Perspectives in Solid Phase Synthesis and Related Technologies: Peptides, Polypeptides and Oligonucleotides*, 1992, pp. 305-309, Intercept, Andover.

C. Kaletta, K.-D. Entian & G. Jung (1991), Eur. J. Biochem. **199**, 411-415.

S. Freund, G. Jung, O. Gutbrod, G. Folkers, W.A. Gibbons, H. Allgaier & R. Werner (1991), Biopolymers **31**, 803-811.

G. Jung (1991), Angew. Chem. **103**, 1067-1084; Angew. Chem. Int. Ed. Engl **30**, 1051-1068.

M. van de Kamp, H.W. van de Hooven, R.N.H. Konings, C.W. Hilbers & F.J.M. van de Ven (1992), Proceedings of the XVth ICMRBS, Jerusalem, p. 192.

G. Bierbaum & H.-G. Sahl (1993), Zbl. Bakt. Hyg. **278**. 1-22.

N. Zimmermann, S. Freund, A. Fredenhagen, G. Jung (1993), Eur. J. Biochem. **216**, 419-428.

G. Bierbaum, J. Süling, C. Szekat & H.-G. Sahl (submitted, FEMS Lett.), *Lysis of Staphylococcus epidermis with the lantibiotic nisin for isolation of DNA and RNA.*
J.C. Scott, H.-G. Sahl, A. Carne & J.R. Tagg (1992), FEMS Lett. **93**, 97-102.
H.-G. Sahl (1992), *Lantibiotics, an introduction,* in: R. James, C. Lazdunski, F. Pattus (eds.), *Bacterial plasimid-coded toxins: bacteriocins, microcins and lantibiotics*, Springer, Heldelberg, pp. 71-74.
H.-G. Sahl (1992) in: R. James, C. Lazdunski, F. Pattus (eds.), *Bacterial plasimid-coded toxins: bacteriomicins and lantibiotics,* Springer, Heidelberg, pp. 93-106.
O.P. Kuipers, H.S. Rollema, W.M.G.J. Yap, H.J. Boot, R.J. Siezen & W.M. de Vos (1992), J. Biol. Chem. **267**, 24340-24346.
W.M. de Vos (1992), M.R. Ladisch, A. Bose (eds.), *Harnessing Biotechnology for the 21st Century*, Am. Chem. Soc., Washington, pp. 524-527.
W.M. de Vos, J.W.M. Mulders, R.J. Siezen, J. Hugenholtz & O.P. Kuipers (1993), *Applied and Environmental Microbiology* **59**, 213-218.
D.H. Vuijk, J.J. Dekkers, H.C. v.d. Plas (eds.), *Developping Agricultural Biotechnology in The Netherlands* (1993); W.M. de Vos, pp. 266-274; O.P. Kuipers, H.S. Rollema, J. Hugenholz, W.M. de Vos, R.J. Siezen, pp. 236-241.
O.P. Kuipers, MM. Beerthuyzen, R.J. Siezen, W.M. de Vos (1993) Eur. J. Biochem. **216**, 281-291.
O.P. Kuipers, H.S. Rollema, W.M. de Vos, R.J. Siezen (1993), FEBS Lett., **330**, 23-27.
O.P. Kuipers, H.S. Rollema, MM. Beerthuyzen, J.R. v.d. Meer, W.M. de Vos, R.J. Siezen (1993), in: E.L. Griffin, H.B. Möllby, C.G. Heden (eds.), *The Lactic Acid Bacteria*, pp. 9-13.
O.P. Kuipers, J. Hugenholtz (1993), De Ware(n) Chemicus **23**, 191-197.
O.P. Kuipers, H.S. Rollema, M.M. Berthuyzen, J.R. v.d. Meer, W.M. de Vos, R.J. Siezen (1993), in: L. Alberghina, L. Frontali, P. Sensi (eds.), *ECB6, proceedings of the 6th European Congress of Biotechnology*, pp. 221-224.
J.R. v.d. Meer, J. Polman, MM. Beerthuyzen, R.J. Siezen, O.P. Kuipers, W.M. de Vos (1993), J. Bacteriol. **175**, 2578-2588.
J.-C. Piard, O.P. Kuipers, H.S. Rollema, M.J. Desmazeaud, W.M. de Vos (1993), J. Biol. Chem. **268**, 16361-16368.
P.J.G. Rauch, Thesis, Agricultural University of Wageningen (1993).
A.G. Beck-Sickinger, G. Jung (1993), Liebigs Ann. Chem. **1993**, 1125-1131.
J.H. Spee, W.M. de Vos & O.P. Kuipers (1993), Nuc. Ac. Res. **21**, 777-778.
W.M. de Vos in: D. Hoover, L. Steenson (eds.), *Bacteriocins of Lactic Acid Bacteria*, Academic Press, New York (in press).
S. Stoffels, H.-G. Sahl, A. Gutmundsdottir (1993), Int. J. Food Microbiol. **20**, 199-210.
R. Jack, A. Carne, J. Metzger, S. Freund, H.-G. Sahl, G. Jung, J.R. Tagg (1994), Eur. J. Biochem. **219**, 699-705.
H.-G. Sahl (1994), Appl. Env. Bactoriol. **60**, 752-755.
J.R. v.d. Meer, H.S. Rollema, R.J. Siezen, O.P. Kuipers, W.M. de Vos (1994), J. Biol. Chem. **269**. 3555-3562.
P.J.G. Rauch, O.P. Kuipers, R.J. Siezen & W.M. de Vos (1994), in: L. de Vuyst, E.J. Vandamme (ed.), *Bacteriocins of Lactic Acid Bacteria: Microbiology, Genetics and Applications*, Elsevier, London, pp. 223-249.
H.-G. Sahl in: H.G. Boman (eds.), *Antimicrobial Peptides*, Wiley (1994), pp. 27-53.
M. Reis, M. Eschbach-Bludau, M.I. Iglesias-Wind, T. Kupke, H.-G. Sahl, Appl. Environ, Microbiol. (in press).
H.E. Meyer, M. Heber, B. Eisermann, H. Korte, J.W. Metzger, G. Jung (1994), Anal Biochem. (in press).
G. Bierbaum, M. Reis, S. Szekat, H.-G. Sahl, Appl. Environ. Microbiol. (submitted).
H.-G. Sahl, R. Jack, G. Bierbaum, Eur. J. Biochem. (submitted, invited review).
C. Meyer, H.-G. Sahl, *Sequence of Pep5 gene cluster* .(in preparation).
G. Bierbaum,M. Reis, H.-G. Sahl, Bio Essays (invited review, in preparation).
H. v.d. Hooven, M. v.d. Kamp, R. Konings, C. Hilbers, F. v.d. Ven, *Structure of nisin bound to detergent micelles* (in preparation).
H. v.d. Hooven, M. v.d. Kamp, R. Konings, C. Hilbers, F. v.d. Ven, *Interaction of nisin with the surfaces of detergent micelles* (in preparation).

Multidisciplinary approach to the analysis of enzyme catalysis, protein stability and folding (BIOT CT-910270)

COORDINATOR:
O. MISSET, Gist-brocades, Delft, NL
PARTICIPANTS:
A.R. FERSHT, University of Cambridge, Cambridge, UK
S. WODAK, Université Libre de Bruxelles, Brussels, B
P.L. MATEO, University of Granada, Granada, E

BACKGROUND INFORMATION

Understanding the forces that determine the stability of native proteins in aqueous solution and the mechanisms that underlie catalysis and the folding/unfolding transition of the polypeptide is essential to the interpretation of biological phenomena at the molecular level and to the ability to design improved proteins. The present project aimed at improving our understanding of these processes by applying a multidisciplinary approach to Barnase, a ribonuclease from *Bacillus amylolique faciens*, ideally suited as a paradigm for this approach.

OBJECTIVES AND PRIMARY APPROACHES

The various disciplines involved experimental techniques such as site-directed mutagenesis, thermal analysis, denaturation studies, enzyme kinetics, X-ray diffraction and 2D-NMR techniques and theoretical methods such as molecular mechanics, knowledge based modelling with dynamical simulations and free energy calculations.

RESULTS AND DISCUSSION

A. Development of the research toolbox

— The following structures have been determined: Barnase (NMR-solution structure) and X-ray structures of the mutants S91A, I96V, I76A, I88A, I96A, I76A + I88A + I96A, Y78F and S91A and wild type Barnase in complex with the tetranucucleotide d(CGAC).
— For the simulations of the free energy change upon histidine deprotonation, energy parameters were estimated.
— New energy parameters were incorporated into the software packages BRUGEL and DISCOVER (BIOSYM) in order to carry out the molecular dynamics simulations of Barnase in the presence of 5 phosphate ions.
— A new Molecular Dynamics technique called *Multiple Copy Simultaneous Dynamics* which allows the efficient simulations of interactions between an enzyme and its substrate (or inhibitor) was implemented into the software package BRUGEL.
— A new (Russian) scanning calorimeter was installed in Granada. The reconstruction of the multi-frequency calorimeter was finalized. Software necessary for the analysis of scanning calorimetry records in order to calculate ligand binding parameters was further developed.
— New algorithms were developed for the determination of the binding parameters of the protein-ligand interaction from the Differential Scanning Calorimetry records.

B. Results aiming at the understanding of the mechanism of folding and stability

*1. **Folding of barnase fragments*** — Fragments of Barnase were (chemically) synthesized in order to study the formation of the tertiary structure of Barnase, via the secondary structure of these fragments. It was found that the major α-helix is formed even as short sequence and does not aggregate in solution. The central β-sheet, however, does tend to form large aggregates.

*2. **Engeneering of disulphide bonds into barnase*** — Disulphide bridges were introduced into Barnase in order to act as probes of the folding pathway of this protein. One disulphide, between residue 85 and 102, links two loops known to pack together early in the folding pathway. A second one, between residue 43 and 80, links two elements of secondary structure known to pack together only **after** the rate determining step of folding.
Both disulphide bridges confer stability (measured by equilibrium denaturation). The kinetics of unfolding and refolding of the mutant proteins have been measured, and a comparison of the disulphide proteins and their corresponding dithiol forms has been made by use of thermodynamic cycles. The disulphide bond engineered into the part of the protein that folds early confers stability upon the intermediate and transition states of folding. The protein with a disulphide bond connecting parts of the protein that fold late is not stabilized until the protein reaches its final folded form. Bridge 43-80 unfolds 19 times more slowly than wild type Barnase and 170 times more slowly than its dithiol form.
NMR H/D exchange experiments have been used to examine the effect of the disulphide bonds on the dynamic behaviour of the mutant proteins. Comparisons of the exchange rates of amide protons in wild-type and the mutant proteins have been made under the same conditions.
A *qualitative* comparison of 3 similar proteins allows you to:

i) identify elements of secondary structure which exchange through local breathing movements and those which exchange only when the protein unfolds considerably.

ii) identify regions which are locally stabilised or destabilised by a mutation.
A *quantitative* comparison of wild type with a single mutant using

$$\Delta\Delta G = - RT^{\ln} \frac{k_{ex}}{k'_{ex}}$$

allows us to distinguish further

i) those protons which exchange by local breathing ($\Delta\Delta G_{ex} \approx 0$),

ii) those protons which exchange by considerable unfolding ($\Delta\Delta G_{ex} \approx \Delta\Delta G_{U\text{-}F}$ 'global infolding mechanism') and

iii) those protons which exchange by both mechanisms ($0 < \Delta\Delta G_{ex} < \Delta\Delta G_{U\text{-}F}$). Furthermore, it allows us to compare equilibrium exchange with protein folding pathways.

Specifically, for these disulphide mutants of Barnase we have been able to show that the disulphide bonds have a local effect on the dynamic behaviour of the protein which can be correlated to the relative stabilising effect of the disulphide bond. We have also shown that there is no direct relationship between exchange behaviour at equilibrium and protein folding pathways.

*3. **Unfolding of basttar, barnase and their complex*** — High sensitivity Differential Scanning Calorimetry (DSC) has been used to characterize the thermal unfolding of the wild type Barnase and Barstar mutants as well as their complex. The necessity of using the mutants of Barstar in DSC studies was caused by the complicated unfolding pattern of the wild-type protein arising from uncontrollable

oxidation of its two sulfhydryl groups upon heating. The barstar mutant C40A + C82A showed a simple two-state unfolding upon heating. At neutral pH, however, its thermal stability was much higher than that of Barnase, despite the fact that its unfolding enthalpy (even when calculated per amino acid) was much lower. At acidic pH, i.e. 2.5-3, heating of the protein inside the calorimeter was not accompanied by a typical endothermic heat effect, indicating that at these conditions Barstar looses its native conformation. It was also found that this 'denaturedprotein has a strong tendency to form soluble oligomers which might account for stabilization of the 'molten globular' state reported in the literature for these solvent conditions. According to DSC data, complexing of Barnase with the Barstar mutant increases the thermal stability of the former by as much as 20°C and that of Barstar by only a few degrees. The thermodynamic parameters were determined both by DSC and IMC (isothermal mixing calorimetry). Their analysis has shown that the relatively high Gibbs energy gain (about -70 kcal/mole) is mostly produced by the negative heat effect of about 60 kcal/mole, while the entropy change at complex formation is close to zero. The results of the calorimetric study on Barstar and its complex with Barnase will be published in another joint paper which is currently in preparation [3].

4. Molecular dynamics of solvated barnase — A 250 picosecond trajectory of solvated Barnase has been generated by molecular dynamics simulations. The system included all the protein heavy atoms, as well as all hydrogen positions generated using standard bond distances and angles, 94 crystallographic positioned water molecules and 2265 generated water molecules.

The simulated protein displayed very modest deviations relative to the crystal structure (1.22 and 1.65 Å rms of the backbone and all atoms respectively). The accessible surface and corresponding accessible volume increased 17% and 8% respectively over the 250 ps time scale, consistent with a slight expansion of the protein (0.5Å increase in gyration radius). A detailed monitoring of the H-bonds formed during the simulations indicated that a large majority of those identified crystallographically are maintained during the simulations. It also revealed frequent exchange between water and protein H-bonds, mostly in loop regions.
The structural and dynamic properties of the water molecules near the protein surface were analyzed by computing radial distribution functions for the water oxygens and hydrogens. The results were in excellent agreement with expected H-bonding distances and with distances between water and non-polar groups obtained in simulations of small water molecules. Self diffusion coefficients of water molecules around the backbone amide protons, the Hß hydrogens of alanines, the H_β hydrogens of isoleucine and the H_ζ hydrogens of lysines in Barnase were calculated. This coefficient is very small ($D = 0.3\text{-}0.4\ 10^{-9}\ m^2.s^{-1}$) relative to the value in bulk ($D = 4.6\ 10^{-9}\ m^2.s^{-1}$) for the water surrounding the amide hydrogens and the $AlaH_\beta$, whereas water in contact with $IleH_\gamma$ and $LysH_\zeta$ hydrogens is significantly more mobile ($D = 1.6\ 10^{-9}\ m^2.s^{-1}$) but still less mobile than bulk water. These results indicate that there is no detectable difference in the dynamic behaviour of water molecules surrounding hydrophobic and polar groups but that this behaviour seems to be influenced by their proximity to the protein backbone.

5. Simulation of the thermal denaturation of the β-hairpin comprising residues 85-102 of barnase — Molecular dynamics simulations of the thermal denaturation of the ß-hairpin comprising residues 85-102 of Barnase were undertaken as part of a larger analysis if the unfolding and stability properties of Barnase. This hairpin was chosen because it adopts its native structure early in the folding process which suggests that it may be relatively stable on its own. This 18-residue segment was

simulated for 300 ps in vacuum and in water respectively and at 3 temperatures: 300K, 450K and 600K using the program CHARMM.
In the room temperature simulations in water, the H-bond pattern of the starting crystal structure is well conserved; a total of 3 classes of conformations is observed of which one is massively populated (95% of the total). This highly populated class is the same as the most populated conformation in the vacuum simulations of the same system where a wider conformational diversity is observed (22 clusters). 43 different classes of conformations are identified in the 450K simulations in aqueous solution and 2 of these classes are also present at room temperature (one of them represents 30% of the structures). During this trajectory, H-bonds at the stem of the hairpin, between residues 85--102 and 87--99 are lost and the C- and N-termini of the hairpin move apart while the rest of the stucture is conserved including the positivie Φ angle of Trp94. Concomitant with these changes, the interactions between hydrophobic side chains near the ß-turn seem to be reinforced as these side chains move closer to one another. This seems to delay the transition towards the completely extended conformation which becomes significantly populated only during the 600K simulation.

C. Results aiming at the understanding of enzyme kinetics

1. The interaction between barnase and barstar — The dissociation constant ($K_D < 10^{-13}$ M) for the Barnase:Barstar complex was determined from experiments measuring the on-rate ($\approx 5*10^8$ $M^{-1}.s^{-1}$) by fluorescence change of Barnase and the off-rate ($10^{-5}s^{-1}$) using tritiated Barstar.

2. The binding of 3'GMP to barnase — Isothermal Mixing Calorimetry revealed that Barnase binds 3'-GMP to a single, high affinity site with a heat effect of about 50 kcal/mole and a heat capacity change of -250 J/K*mole. Both free Barnase and Barnase in complex with 3'-GMP undergo a highly reversible, two-state thermal unfolding process in the pH-range 2-5 with the increments of the free energy and heat capacity similar to those of other small globular proteins. However, both the enthalpy and entropy of unfolding turn out to be unusually high at the denaturation temperature of Barnase. It was also found unexpectedly that the thermodynamic unfolding properties of Barnase do not fit into correlations between ΔH and ΔC_p described in the literature for small globular proteins. This discrepancy might be related to particular features of the folded and/or unfolded states of the protein. See joint publication [1].

3. Modelling of the barnase: pentanuclaotide transition state model — The amended Barnase:pentanucleotide transition state model differs from the intial proposal in the details of the interactions made between the enzyme and the nucleotide. Using the detailed three dimensional picture of the Barnase:transition state complex, experimental observations on the kinetics of oligonucleotide catalysis in Barnase could be rationalized.

4. Molecular dynamics simulation of barnase in the preserce of 5 phosphate ions and water — This study was carried out with the purpose to get more information about substrate recognition and binding. The system consisted of one Barnase, 5 methylphosphate and 2359 water molecules. A rectangular box of 82681Å3 was used for the simulation under periodic boundary conditions and at 300 K with constant box volume (see joint publication [2]

5. Simulation of histidine (de)protonation — The approach to the problem involved computing the free energy changes associated with the (de)protonation of His18 in the native protein (ΔG_P), and with a histidine dipeptide unit in water (ΔG_W). The pK_a-shift was calculated from the equation $\Delta pK_a = (\Delta G_P - \Delta G_W)/(2.3RT)$.

The result was $\Delta pK_a = 4.0 \pm 1.0$. Although qualitatively in good agreement with the experimental value ($\Delta pK_a = 1.6$ units), it is still a factor 2 too high, probably due to errors in treating electrostatic interactions, caused by applying a distance cut-off in computing non-bonded interactions. Also the from fluorescence experiments determined ΔpK_a of 0.6 pH-units, caused by the aromatic-histidine interaction between Trp94 and His18, could reasonably be substantiated by calculations: 0.40 pH-units.

The finite difference solution to the Poisson-Boltzmann equation, implemented in the program *Delphi*, was used to evaluate the shifts in the pK_a of catalytic residues in Barnase produced by the binding of oligonucleotide ligands. Calculations of the pK_a shifts were performed using Barnase-oligonucleotide complexes modelled in Brussels and the crystal structure of Barnase-d(CpGpApC) determined in Cambridge. The results show that the binding of a ribonucleotide substrate induces significant shifts in the pK_a's of Glu73 and His102. Based on the hypothesis that the titration profiles of these two residues determine the pH profile of the catalytic reaction in Barnase, it is shown that binding of nucleotides produces an upward shift in the pH-optimum for catalysis by about 3.1 units relative to the pH optimum of 5.25 calculated considering experimental pK_a values in presence of bound 3'-GMP. This shift is in very good agreement with the pH optimum of 8.5 determined experimentally. The calculations show that, in both the modelled and the crystallographic complexes, the nucleotide groups contributing most to this shift are (in order of decreasing influence): the +2 and 0 phosphates, the +1 adenine and +1 ribose moieties. The main features responsible for these contributions, and common to both complexes, are the stacking of the +1 adenine base onto the imidazole of His102 and the binding mode of the +2 phosphate.

MAJOR SCIENTIFIC BREAKTHROUGHS

i) Solution structure of Barnase by NMR;
ii) A detailed 3D-model for the Barnase: pentanucleotide transition state complex;
iii) By determining which regions in a protein unfold early, stabilization may be achieved by introduction of disulphide crosslinks.

MAJOR COOPERATIVE LINKS

The Cambridge group (8-10 scientists) carried out most of the experimental techniques. Thermal analyses were carried out by (4) who was supplied by (2) with pure Barnase and Barstar. (3) collaborated with (1) on the molecular dynamics simulations of Barnase and compared together with (2) NMR- and X-ray structural results with theoretical calculations.
Eight joint meetings took place in Delft, Cambridge, Brussels and Granada.

PUBLICATIONS

Joint publications

Martinez, J.C., Harrou, M.E., Filimonov, V.V., Mateo, P.L. and Fersht, A.R. (1994), *Biochemistry*, **33**, 3919-3926 *'A Calorimetric Study Of The Stability Of Barnase And Its Interaction With 3'GMP'*

Aehle, W., Misset, O. and Belle, D. van (1992), Abstract 3rd Nordic Conference on Protein Engineering, August 12-16, 1992 *'Molecular Dynamics Simulation of Barnase in water in presence of a substrate analogue'*

Martinez, J.C., Filimonov, V.V., Mateo, P.L., Schreiber, G. and Fersht, A.R. *A Calorimetric Study Of Barstar and Its Complex With Barnase;* in preparation

Individual publications

Buckle, A.M. and Fersht, A.R. (1994), *Biochemistry,* **33**, 1644-1653.
Chen, Y.W., Fersht, A.R. and Henrick, K. (1993), *J. Mol. Biol.,* **234**, 1158-1170.
Clarke, J. and Fersht, J. (1993), *Biochemistry,* **32**, 4322-4329.
Clarke, J., Hounslow, A.M., Bycroft, M. and Fersht, A.R. (1993), *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9837-9841.
Kippen, A.D., Sancho, J. and Fersht, A.R. (1994), *Biochemistry,* **33**, 3778-3786.
Wodak, S.J., van Belle, D. and Prévost, M. (1994) in *Molecular Simulations*, Julia Goodfellow edt. (in press).
Lippens, G., van Belle, D., Wodak, S.J. and Jeener, J. (1993), *Molecular Physics,* **80**, 15469-1484.
Wodak, S.J., Van Belle, D. and Prévost, M. (1992) in *Proceedings of INSERM workshop 49 'Molecular modelling and graphics applied to structure-function relationships* (Le Vésinet, France, November 1992).
Horovitz, A. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 773-740.
Sancho, J. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 741-747.
Sancho, J., Neira, J.L. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 749-758.
Loewenthal, R., Sancho, J. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 759-770.
Fersht, A.R., Matouschek, A. and Serrano, L. (1992), *J. Mol. Biol.* **224**, 771-782.
Serrano, L., Kellis, J.T.Jr., Cann, P., Matouschek, A. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 783-804.
Serrano, L., Matouscheck, A. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 805-818.
Matouschek, A., Serrano, L. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 819-835.
Matouschek, A., Serrano, L. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 837-845.
Serrano, L., Matouschek, A. and Fersht, A.R. (1992), *J. Mol. Biol.* **224**, 847-859.
Willaert, J. *et al.*, (1992) *Biochemistry,* **31**, 711-716.
Sancho, J., Serrano, L. and Fersht, A.R. (1992), *Biochemistry,* **31**, 2253-2258.
Serrano, L., Neira, J.L., Sancho, J. and Fersht, A.R. (1992), *Nature,* **356**, 453-455.
Meiering, E.M., Serrano, L. and Fersht, A.R. (1992), *J. Mol. Biol.* **225**, 585-589.
Serrano, L., Sancho, J., Hirshberg, M. and Fersht, A.R. (1992), *J. Mol. Biol.* **227**, 544-559.
Horovitz, A., Matthews, J.M. and Fersht, A.R. (1992), *J. Mol. Biol.* **227**, 560-568.
Serrano, L., Bycroft, M. and Fersht, A.R. (1991), *Pure and Applied Vhemistrty,* **63**, 187-194.
Horovitz, A., Serrano, L. and Fersht, A.R. (1991), *J. Mol. Biol.,* **219**, 5-9.
Sali, D., Bycroft, M. and Fersht, A.R. (1991), *Techniques in Protein Chemistry II,* **28**, 295-303.
Fersht, A.R., Bycroft, M., Horowitz, A., Kellis Jr., J.T., Matouschek, A. and Serrano, L. (1991), *Phil. Trans. R. Soc. Lon. B.,* **332**, 171-176.
Loewenthal, R., Sancho, J. and Fersht, A.R., *Biochemistry,* **30**, 6775-6779.
Sali, D., Bycroft, M. and Fersht, A.R. (1991), *J. Mol. Biol.,* **220**, 779-788.
Dancho, J., Meiering, E.M. and Fersht, A.R. (1991), *J. Mol. Biol.* **221**, 1007-1014.

Stability studies and protein-design studies with triosephosphate isomerases (BIOT CT-900182]

COORDINATOR:
A) R.K. WIERENGA, EMBL, Heidelberg, D
PARTICIPANTS:
B) J. MARTIAL, Univ. Liège, Liège, B
C) F.R. OPPERDOES, ICP, Brussels, B
D) W.G.J. HOL, Univ. Groningen, Groningen, NL (until 1.9.1992).
E) H. KESSLER, Techn. Univ. Münich, Münich, D
F) A. RENARD, Eurogentec SA, Liège, B

BACKGROUND INFORMATION

A network of European research groups has been formed to study and engineer the stability and activity of the enzyme triosephosphate isomerase (TIM). TIM is a dimeric glycolytic enzyme. Each subunit has the TIM-barrel fold. The TIM-barrel fold is observed in many different enzymes catalysing very diferent reactions. Clearly this frame work is able to create very different binding pockets. The ultimate goal of this project is to create a new binding site on a TIM-barrel framework. The available expertise in this network concerns molecular biology and biochemistry (A,B,C), crystallography and modelling (A,D), cyclopeptide synthesis (E) and mass production of over expressed proteins (F).

OBJECTIVES AND PRIMARY APPROACHES

The initial objectives can be summarised as follows:

1) Characterisation of extremophylic TIM's in order to learn more about thermostability.
2) Mutagenesis of human TIM to make it more stable, mutagenesis of trypanosomal TIM in order to obtain a monomeric variant.
3) Binding studies with cyclohexapeptides designed to influence the properties of wildtype TIM or its variants.

RESULTS AND DISCUSSION

Characterisation of psychrophylic and thermophylic TIM (Martial, Hol).
Four TIM sequences of extremophylic bacteria were determined (Table 1). These sequences show considerable sequence variety. Therefore it is not trivial to establish which factors are responsible for the different thermal stabilities. For the two moderate extremophyles, *Moraxella* spp and *Bacillus stearothermophilus*, a detailed comparison has been carried out. From this comparison it appears that the adaptability of TIM to high temperature is favoured by better stabilising residues for the helix dipole as well as better helix-forming residues whereas the adaptability of TIM to low temperature seem to reside in the nature of helix capping residues.

Over expression and purification to homogenity of cold and hot TIM has been done. Hot TIM could be crystallised and a 2.8Å refined crystal structure of hot TIM has become availabe recently. The *in vitro* stability of hot and cold TIM has been quantified. It has been shown that the enzymatic activities of the cold and hot TIM are strongly dependent on the incubation temperature used. While, under the given experimental conditions cold TIM is rapidly inactivated at temperatures above 35°C, hot TIM is stable at temperatures up to 65°C. Thus, compared to the inactivation temperature, around 50°C of the mesophylic E.coli TIM, the cold TIM

presents indeed psychrophylic characteristics, whereas hot TIM can be considered as thermophylic.

More recently overexpression experiments with the genes of supercold and superhot TIM's, show that it will be possible to obtain pure supercold and superhot TIM. Subsequently the biochemical properties, such as stability will be measured. In addition it will be attempted to grow crystals of these supercold and superhot TIM's.

Table 1: TIM-sequence data of TIM from four different extremophylic bacteria

			sequence identity (%)			
		No. of residues	*super cold*	*cold*	*hot*	*super hot*
supercold	*Vibrio marinus*	256	–	39	41	41
cold	*Moraxella sp* TA137	270	–	–	34	34
hot	*Bacillus stearo-thermophilus*	253	–	–	–	51
superhot	*Thermatoga maritimia*	255	–	–	–	–

More stable variants of human TIM
(Martial, Hol).

Following up on the sequence analysis of TIM's with different thermal stability it was decided to make human TIM variants with predicted increased thermal stability by choosing mutations which would stabilise α-helices. Three approaches were considered, involving the replacement of solvent exposed residues in α-helices with alanines, the correct positioning of prolines at the N-cap position of α-helices and the stabilisation of the helix dipoles by replacing the N-terminal residues of some helices with negatively charged residues (Asp). Two mutants were already constructed, where Gln179 was replaced with an Ala or an Asp. The Q179D replacement significantly stabilised human TIM, while the Q179A replacement did not affect the protein stability. A third mutant, A215P, has now been constructed; the introduction of this proline at the N-cap position of α7 increased the stability of human TIM. The construction of other two mutants is now in progress: K193A and S105D.

Point mutation variants of trypanosomal TIM and human TIM
(Wierenga, Martial, Opperdoes).

Three point mutation variants of interface residues of trypanosomal TIM have been characterised: C14G, H47N and V46G. These variants are monomeric at low protein concentration ($\leq$1mg/ml) and dimeric at high protein concentration ($\geq$20mg/ml). The monomeric forms are not molten globule like proteins, but instead they are compactly folded and the residual active site is still able to bind 2-phosphoglycolate. The dimeric forms have retained most of the catalytic efficiency. The monomeric forms are considerable less stable than the wild type dimer.

Two point mutation variants of interface residues of human TIM have been characterised: M14Q and R98Q. These point mutations variants have similar properties as C14G, H47N and V46G. Interestingly R98Q is completely inactive. Arg98 is a completely conserved residue in the helical fragment of loop-4. The complete inac-

tivity of R98Q points to an important function of Arg98 for achieving the proper active site architecture.

Binding studies with cyclohexapeptides
(Kessler, Opperdoes, Wierenga)
At the beginning of this project two classes of cyclohexapeptides of interest to the TIM project were available. These two categories of peptides were choosen for further study because of the general interest in the mode of binding between cyclopeptides and proteins.

Class II cyclopeptides are rather hydrophobic compounds, with inhibitory properties against trypanosomal TIM, but not against other TIM's. On further investigation it turned out that these inhibitory properties were due to coprecipitation of trypanosomal TIM with the hydrophobic cyclopeptides in the assay mixture, and not due to a specific complex formation between cyclopeptide and trypanosomal TIM. Therefore studies at the molecular level of the mode of binding of these cyclopeptides are not possible.

Class I cyclopeptides are analogs of loop-3 (the major interface loop) and therefore pontential inhibitors of dimerisation. On testing it turned out that they are not inhibitors of TIM. The superposition of the crystal structure of one of these cyclohexapeptides on loop-3 in wildtype TIM suggested, that the conformation of this cyclopeptide would create a clash between the side chain of Cys13 and the cyclohexapeptide. Therefore the effects of this cyclohexapeptide were also tested on the pointmutation variant C14G, but no effect was measured.

Apparently these cyclohexapeptides are too rigid and/or too small to bind with measurable affinity in the binding pocket of loop-3. This approach will be pursued further by testing the binding properties of much longer, but linear peptide analogues of loop-3, for monoTIM and the point mutation variants.

Octarellin: A project concerning the *de novo* protein design of a TIM-barrel protein.
(Martial)
Originally octarellin-1 has been characterised. Octarellin-1 has eightfold symmetry. Current studies focus on the characterisation of octarellin-III, which is based on four fold symmetry. The facilities of Eurogentec have been used for the overexpression in a 25l fermentor and purification of octarellin.

***Leishmania mexicana* TIM**
(Opperdoes, Wierenga)
In *Leishmania* only one TIM gene is present, although TIM occurs both in the cytoplasm as well as in the glycosome. The residues constituting the subunit interface are highly conserved between the enzyme of *L. mexicana* and *T. brucei*, but are mostly different from those TIM's of other organisms. One major substitution is present in the interface region of the *L. mexicana* protein: a glutamate at positon 65, instead of glutamine in all other known 30 triosephosphate isomerase sequences. The glutamine is supposed to be important for the stability of the dimeric enzyme.

L. mexicana triosephosphate isomerase has been overexpressed in *Escherichia coli* and the enzyme has been purified to near-homogenity. It appears a stable dimeric protein. Cystallisation experiments are being carried out in order to be able to analyse the structural differences near position 65.

Modular mutagenesis of *E. coli* TIM
(Martial, Wierenga)

The crystal structure of a hybrid *E. coli* TIM has been determined at 2.8Å resolution. The hybrid TIM (ETIM8CHI) was constructed by replacing the eigth βα-unit of *E. coli* TIM with the equivalent unit of chicken TIM. This replacement involves 10 sequence changes. One of the changes concerns the mutation of a buried alanine (Ala232 in strand-8) into phenylalanine. The ETIM8CHI-structure shows that the A232F sequence change can be incorporated by a side chain rotation of Phe224 (in helix-7). No cavities or strained dihedrals are observed in the hybrid enzyme, and ETIM8CHI and *E. coli* TIM have similar stabilities. The largest CA-movements, approximately 3Å, are seen for the C-terminal end of helix-8 (associated with the outward rotation of Phe224) and for the residues in the loop after helix-1 (associated with sequence changes in helix-8).The active site architectures of wildtype *E. coli* TIM and ETIM8CHI show no obvious differences, therefore it is not clear why the k_{cat} of ETIM8CHI is 10 times lower than in wild type *E. coli* TIM. Possibly subtle changes in the structural or dynamic properties between wild type *E. coli* TIM and ETIM8CHI, which are not apparent from the comparison of the structures, are responsible for the different kinetic constants.

Characterisation of monoTIM, derived from trypanosomal TIM
(Wierenga)

Protein engineering on trypanosomal TIM converted this oligomeric enzyme into a stable, monomeric protein that is enzymatically active. Wild-type TIM consists of two identical subunits that form a very tight dimer involving interactions of 32 residues of each subunit. By replacing 15 residues of the major interface loop by another 8-residue fragment, a variant, referred to as mono TIM, was constructed that is a stable and monomeric protein, with TIM activity. The length, sequence, and conformation of the designed fragment were suggested by extensive modeling.

The crystal structure of monoTIM shows that it retains the characteristic TIM-barrel (βα)8-fold and that the new loop has a structure very close to that predicted by modelling. Two other interface loops, loop-1 and loop-4, which contain the active site residues Lys13 and His95, respectively, show significant changes in structure in monoTIM compared with dimeric wild-type TIM. MonoTIM has residual catalytic activity. This residual activity is an intrinsic property of monoTIM. The Lys13 is crucial for this activity, because monoTIM (K13A) is inactive and monTIM (K13R) is less active than monoTIM. The residual catalytic activity can be explained by assuming that in solution loop-1 and loop-4 are flexible and in the presence of substrate can adopt the active site geometry required for supporting catalysis. This assumption is supported by several observations. For example, a point mutation derivative of monoTIM, referred to as monoSS, has also been crystallised. MonoSS is obtained from monoTIM by mutating Phe45 and Val46 into serines. In the crystallised monoSS loop-4 has adopted again the wild type conformation, whereas loop-1 has adopted a conformation different from monoTIM and different from wild type TIM. These structural difference do not seem to be related to the two point mutations, but rather due to differences in crystalisation conditions and crystal packing. Despite the fact that the structural differences are effecting the conformation of the catalytic residues, monoTIM and monoSS have the same catalytic properties, in agreement with the assumption that in solution these loops are flexible.

Further protein design studies with monoTIM will now focus on loop changes aimed at making monoTIM as active as wild type TIM.

MAJOR SCIENTIFIC BREAKTHROUGHS

1. The sequencing and characterisation of TIM from *Vibrio marinus, Moraxella sp* TA137, *Bacillus stearothermophilus* and *Thermatoga maritima*. These bacteria are extreme psychrophylic, psychrophylic, thermophylic and extreme thermophylic organisms respectively.

2. The determination of the crystal structures of *E. coli* TIM, human TIM, *B. stearothermophilus* TIM and ETIM8CHI.

3. The structural characterisation of monoTIM, a monomeric variant of trypanosomal TIM.

MAJOR COOPERATIVE LINKS

Each year 2 joint meetings have been organised. At these lively meetings the results obtained by the groups and the joint publications have been discussed intensively.

PUBLICATIONS

Joint publications

Borchert, T.V., Pratt, K., Zeelen, J.Ph., Callens, M., Noble, M.E.M., Opperdoes, F.R., Michels, P.A.M. and Wierenga, R.K. (1993), 'Overexpression of trypanosomal triosephosphate isomerase in *Escherichia coli* and characterisation of a dimer-interface mutant'. Eur. J. Biochem. **211**, 703-710.

Callens, M., Roy, J.V. Zeelen, J.Ph., Borchert, T.V., Nalis, D., Wierenga, R.K. and Opperdoes, F.R. (1993), 'Selective interaction of glycosomal enzymes from trypanosoma brucei with hydrophobic cyclic hexapeptides'. Biochem. Biophys. Res. Commun. **195**, 667-672.

Kessler, V.H., Matter, H., Geyer, A., Diehl, H.-J., Kock, M., Kurz, G., Opperoes, F.R., Callens, M. and Wiereng, R.K. (1992), 'Selektive Inhibierung der trypanosomalen Triosephosphat-Isomerase durch ein Thiopeptid'. Angew. Chem. **104**, 343-345.

Kohl, L., Callens, M., Wierenga, R.K., Opperdoes, F.R. and Michels, P.A.M. (1994), 'Triosephosphate isomerase of *Leishmania mexicana* mexicana Cloning and characterization of the gene, overexpression in Escherichia coli and analysis of the protein'. Eur. J. Biochem. **220**, 331-338.

Mainfroid, V., Goraj, K., Rentier-Delrue, F., Houbrechts, A., Loiseau, A., Gohimont, A.-C., Noble, M.E.M., Borchert, T.V., Wierenga, R.K. and Martial, J.A. (1993), 'Replacing the (βα)-unit 8 of *E. coli* TIM with its chicken homologue leads to a stable and active hybrid enzyme'. Protein Eng. **6**, 893-900.

Mande, S.C., Mainfroid, V., Kalk, K.H., Goraj, K., Martial, J.A. and Hol, W.G.J. (1994), 'Crystal structure of recombinant human triosephosphate isomerase at 2.8 Å resolution. Triosephosphate isomerase-related human genetic disorders and comparison with the trypanosomal enzyme'. Protein Sci. **3**, 810-821.

Rentier-Delrue, F., Mande, S.C., Moyens, S., Terpstra, P., Mainfroid, V., Goraj, K., Lion, M., Hol, W.G.J. and Martial, J.A. (1993), 'Cloning and overexpression of the triosephosphate isomerase genes from psychrophilic and thermophilic bacteria: structural comparison of the predicted protein sequences'. J. Mol. Biol. **229**, 85-93.

Individual publications

Borchert, T.V., Abagyan, R., Jaenicke, R. and Wierenga, R.K. (1994), 'Design, creation, and characterization of a stable, monomeric triosephosphate isomerase'. Proc. Natl. Acad. Sci. U.S.A. **91**, 1515-1518.

Borchert, T.V., Abagyan, R., Radha Kishan, K.V., Zeelen, J.Ph. and Wierenga, R.K. (1993), 'The crystal structure of an engineered monomeric triosephosphate isomerase, monoTIM: the correct modelling of an eight-residue loop'. Structure **1**, 205-213.

Noble, M.E.M., Zeelen, J.Ph. and Wierenga, R.K. (1993), 'Structure of the 'Open' and 'Closed' State of Trypanosomal Triosephosphate Isomerase, as Observed in a New Crystal Form: Implications for the Reaction Mechanism'. Proteins. **16**, 311-326.

Noble, M.E.M., Zeelen, J.Ph. and Wierenga, R.K. (1993), 'Structure of Triosephosphate Isomerase from *Escherichia coli* Determined at 2.6 Å Resolution'. Acta Crystallogr. D**49**, 403-417.

Radha Kishan, K.V., Zeelen, J.Ph., Noble, M.E.M., Borchert, T.V. and Wierenga, R.K. (1994), 'Comparison of the structures and the crystal contacts of trypanosomal triosephosphate isomerase in four different crystal forms'. Protein Sci. 3, 779-787.

Design and engineering of alpha-helical bundle proteins: modified structure and novel functions (BIOT CT-910262)

COORDINATOR:
C. SANDER, EMBL, Heidelberg, D
PARTICIPANTS:
P. ARIOSIO, Univ. of Milano, Milano, I
P.J. ARTYMIUK, Univ. of Sheffield, Dept. Mol. Biol., GB
H. BLÖCKER, GBF, Braunschweig, D
G. CESARENI, Univ. of Roma, Dpto. di Biologia II, Bolognia, I
H.-J. HINZ, Univ. of Münster, Inst. of Phys. Chemistry, Münster, D
M. KOKKINIDIS, IMMB, Dept. of Prot. Structure, Iraklion, GR

BACKGROUND INFORMATION

Protein engineering as part of biology-derived nanotechnology can benefit enormously from detailed understanding of structures at the atomic level. Such understanding becomes the basis of designing proteins as novel materials and carriers of medically and industrially useful molecular function.

OBJECTIVES AND PRIMARY APPROACHES

The goal was to understand, by a combination of theory and experiment, the structural principles of the 4-α-helix bundle proteins *Rop* and ferritin as a basis for engineering novel functions. This interdisciplinary BRIDGE-funded consortium employed techniques of molecular genetics, protein purification, functional and immunoassays, 3-D structure determination (X-ray, NMR), light spectroscopy, microcalorimetry, molecular modelling and protein structure theory.

RESULTS AND DISCUSSION

The work on ferritin structure, assembly and engineering as carrier protein was the result of close collaboration between the Roma, Milano and Sheffield groups. The work on the structure, folding and redesign of *Rop* protein involved the Iraklion, Roma, Münster, Braunschweig and Heidelberg groups.

X-ray structures of ferritin variants (Sheffield)

The work has concentrated on the X-ray analysis of the 4-helix bundle iron-storage protein ferritin, and its comparison with other four-helix bundle structures. Our strategy has been to use site-directed mutagenesis to alter internal residues in order to observe the structural effects. These results are being compared with similar experiments on *Rop*.

Site-directed mutants of human H ferritin

(i) We achieved a 1.9Å crystal structure of the A222 mutant of human H ferritin. In this mutant the ferroxidase site residues are converted into their inactive L equivalents (E62K, H65G) and in addition various metal-binding residues are converted to alanine (E61A, E64A, E67A).

(ii) In the Y34F mutant (2.5Å resolution), Tyr 34 is close to the ferroxidase site and appears to be important for forming an Fe(III)-tyrosinate complex during iron uptake.

Species variants of ferritin.

(i) *Horse L ferritin.* We have solved a 1.95Å structure of recombinant horse spleen L ferritin. Mammalian H and L ferritins have 55% sequence identity and some of the most interesting substitutions lie at the centre of the four helix bundle and involve the replacement of the metal-binding H chain ferroxidase centre residues with salt bridges.

(ii) Schistosoma *ferritin.* We have solved the structure of ferritin from the human parasite *Schistosoma mansoni* (2.6Å resolution). This is important because of the competition for iron between parasite and host in such infections. This protein has 40% identity with the mammalian H and L ferritins.

(iii) E.coli *ferritin.* We have also solved the structure of ferritin from *Escherichia coli* at 2.5Å. This protein is quite distinct from the haem-containing bacterioferritin, or cytochrome b1 (see below). *E.coli* ferritin has only 20% sequence identity with the mammalian and *Schistosoma* ferritins. There are therefore many interesting amino acid substitutions. These include a number of coupled amino acid substitutions within the bundle (example in Fig. 1). Also, *E.coli* ferritin lacks a normally strongly conserved salt-bridge which link the end of the B helix with the beginnings of the A and D helices respectively.

(iv) E.coli *bacterioferritin.* Bacterioferritins are cytochrome b1 molecules with very distant sequence homology (18%) to ferritins. Our 6Å resolution MIR electron density map shows the arrangement of helices in four helix bundles.

Metal-binding sites, comparisons with other four-helix bundle proteins

The *E.coli* ferritin structure solution yielded the first ever complete dataset of an iron-bound ferritin, using the 0.9Å wavelength radiation at Daresbury, revealing three iron sites, two in the ferroxidase centre and one nearby on the inside of the

Coupled amino acid changes in the ferritin core

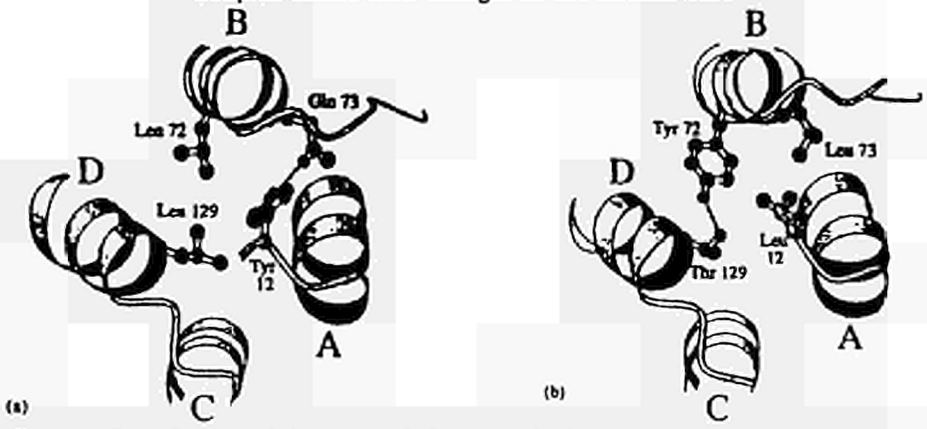


Figure 1: Example of coupled amino acid changes in the ferritin core.
An example of coupled amino acid changes within the ferritin 4-helix bundle found when comparing mammalian ferritins (a), and E.coli *ferritin (b). Tyr 12, Leu 72, Gln 73, Leu 129 in human H ferritin become respectively Leu, Tyr, Leu and Thr; the Tyr12-Gln73 hydrogen bond linking helices A and B disappears but a new Tyr72-Thr129 hydrogen bond appears between helices B and D.*

shell. This has allowed us to make direct comparisons with other iron-binding sites in the four-helix proteins hemerythrin and ribonucleotide reductase.

Ferritin assembly and engineering (Milano)

Ferritin renaturation

Ferritin subunits have an α-helical bundle structure and assemble to form highly stable 24-mer proteins. We studied ferritin refolding and reassembly. Far ultraviolet circular dichroism and fluorescence analyses showed that most ferritins and variants refold in similar conditions, irrespective of their large differences of denaturing conditions. Most mutants readily formed the helical bundle, but reassembly was affected by the alterations of inter-chain interactions (del 1-13 and Leu169->Arg) with the formation of subunit dimers, and by the alteration of the hydrophobic interactions along the two fold axes where two subunits come together, with the formation of folded subunit monomers. H and L subunit variants were renatured together thus obtaining heteropolymers with restricted heterogeneity and the expected subunit composition. When the amino acids near the 2, 3 and 4 fold symmetry axes are modified, refolding mutants accumulated in either monomeric, dimeric or 8-10-meric sub-assembly species. These findings indicate that dimer formation is the first and essential step for ferritin assembly, and that the nature of the interactions along the two fold axes is important for the construction of polyfunctional molecules.

Engineering a ferroxidase centre

A ferroxidase centre was introduced into the human ferritin L chain. The introduction of the four residues which act as metal ligands in the H chain inhibited folding of the L chain *in vivo* and *in vitro*. However, folding to a functional protein with ferroxidase activity occurs in heteropolymers (about 10% L-chain mutant and 90% L-chain wild type, analogous to heteropolymers of 7% H-chain and 90% L-chain). In this way it is possible to construct a ferroxidase centre in the L chain.

Ferritin as a carrier protein

The C-terminal sequence of H ferritin was genetically fused with a 10 Kdal allergen. The construct was efficiently expressed by *E. coli* as a soluble and stable protein which retained most of structural and biochemical properties of ferritin and the allergenic properties of the allergen. It is suggested that fusion to ferritin overcomes the problems of solubility and high levels of expression often found for the production of peptides in *E. coli*.

Ferritin assembly and engineering as a vaccine carrier (Roma)

Engineering a highly immunogenic protein by genetic grafting of the interleukin 1-β peptide VQGEESNDK onto the ferritin structure

Work in the early stages of the project taught us a small number of rules to be followed in modifying the loop connecting helix D and E in ferritin so as not to disturb its the folding and assembly. We thus were able to construct number of hybrid proteins with specific peptides inserted in this region. Particularly interesting is the case of an interleukin 1-β oligopeptide VQGEESNDK that has the immunostimulatory properties of the native protein. In collaboration with a group in Siena we attempted to increase the immunogenicity of ferritin by inserting the peptide either between amino acid Pro161 and Ser162 in the D-E loop (exposed to solvent) or at the carboxy-terminus (inside the protein cage). Both hybrids

assemble correctly into multimeric proteins very similar to wild type. The first protein is recognized, in the native form, by an antibody specific for the VQGEESNDK peptide, while the second is not, consistent with the inside/outside localization. Injecting the first hybrid into mice, the anti-ferritin immune response is increased up to levels comparable to that obtained by coinjection of wild type ferritin and native interleukin 1 . No comparable results are obtained when injecting the carboxy-terminally extended ferritin, suggesting that the immunostimulatory function of the oligopeptide requires its exposure to the solvent on the exterior surface of the ferritin assembly.

Crystal structures of *Rop* variants (Iraklion)

Crystallography

In order to test the sequence requirements for loops in 4-helix bundles, a mutant (RM7-Rop) was designed with an uninterrupted heptad pattern created by the insertion of two Ala residues in the loop. Surprisingly, the mutant protein folds again as a 4-α-helical bundle, albeit with an unusual bend conformation. This confirms that folding in such bundles is dominated by helical packing interactions and is tolerant to changes in the loop; loops adjust to the structural framework of the bundle and can even accomodate residues in unfavourable conformations. The structure of RM7-Rop at the unusually high resolution of 1.07Å — reveals many exciting details. Anisotropic temperature factors, together with standard deviations for atomic positions, were determined for the first time and allow the resolution of multiple side chain conformations and of the rigid body motions of the alpha helices. Another mutant, RM6-Rop (del 30-34) is a very thermostable protein (Tm approx. 100 °C) that crystallizes in space group C2. The other bend mutant, A31P-Rop, crystallizes in three different forms. There are indications that the protein has considerable conformational changes. Work on these structures continues.

Computations and design

As a basis for the prediction and rational design of 4-alpha-helical bundles, statistics on position-specific amino acid preferences were derived and a probability-based local alignment technique developed. The alignment technique is general enough to be used with any sequence and has been implemented in the GCG package. Analysis of *Rop* mutants with 'designed' cavities in the core suggest that the reduction of protein stability due to the cavity is correlated a) with the size of the cavity and b) with the change in the packing efficiency of the core. Also, a *Rop* mutant (called FREAK) with an insertion of a hemerythrin loop in helix 2 was designed and mutagenesis work started. We predict that in this particular case the monomer will fold as a 4-helix bundle.

Wet chemistry — novel techniques

Bend mutants of *Rop* or unstable mutants are very difficult to crystallize. The crystallization of proteins was approached in a more systematic way by studying their solubilities in two-component precipitating systems. This led to the determination of solubility curves which reveal universal properties of proteins in these systems and allow rational determination of crystallization conditions.

Thermodynamic stability and folding of *Rop* variants (Münster)

Folding of native Rop

Detailed thermodynamic and spectroscopic studies were carried out on *Rop* wildtype and mutated proteins to establish a quantitative basis for the contribution of noncovalent interactions to the stability of 4-helix-bundle proteins. The energetics of both heat- and GdnHCl-induced denaturation were measured by differential scanning microcalorimetry (DSC) and/or by following the change in circular dichroism in the far UV-range. Sedimentation equilibrium analyses were performed to characterize the state of aggregation of the protein. No intermediate species could be detected during thermal unfolding of the dimer in the absence of GdnHCl. Under these conditions *Rop* unfolding exhibits a strict two-state behaviour. The thermodynamic parameters for the reaction $N_2 \rightleftharpoons 2D$ are ΔH_D = 580 ± 20 kJ·(mole of dimer)$^{-1}$, ΔC_p = 10.3 ± 1.3 kJ·(mole of dimer)$^{-1}$·K^{-1}, and T_m = 71.0 ± 0.5°C. The corresponding Gibbs energy change of unfolding is $\Delta G_D{}^0$ = +71.7 kJ·(mole of dimer)$^{-1}$ at 25°C and pH 6. In the presence of 2.5 M GdnHCl, however, *Rop* dissociates into monomers at elevated temperatures, as the loss of the concentration dependence of T_m and the decreased molecular weight demonstrate. The corresponding transition parameters are ΔH_D (2.5 M GdnHCl) = 130 ± 10 kJ·(mole of monomer)$^{-1}$ and T_m = 51.6 ± 0.3°C. Isothermal unfolding studies at 19°C using GdnHCl as denaturant yielded a Gibbs energy change of unfolding of 22.4 kJ·(mole of monomer)$^{-1}$. This extrapolated value is 38% lower than the corresponding $\Delta G_D{}^0$-value of 35.85 kJ·(mole of monomer)$^{-1}$ calculated from thermal unfolding for the monomer in the absence of GdnHCl where the protein is known to be a dimer. The result suggests that subunit interactions are an important source of stabilization of the native four-helix-bundle structure of Rop.

Folding of Rop variants

The equilibrium stability properties of a series of core mutants of *Rop* have revealed interesting effects. The L48A and L41A exchanges result in significantly different perturbations of the overall stability of the protein despite the apparent structural identity of the point mutations. Similar complications arose in the kinetic studies dealing with the time course of the unfolding and refolding reactions of wild type and mutated proteins. Thus, more work is required to understand the forces that underly mutational energetics both with regard to equilibrium and activation parameters.

Redesigning *Rop* topology, loops and core (Braunschweig, Heidelberg)

The *Rop* protein was redesigned by altering the topology of loop connections and turning a dimer into a monomer. For the left handed monomer (LmRop, Fig. 2), X-ray data of crystals and results of NMR spectroscopy have confirmed the stability and correct fold of the reengineered protein. Right handed topology variants (RmRop) have been expressed and crystallized, but NMR analysis was not (yet) successful.

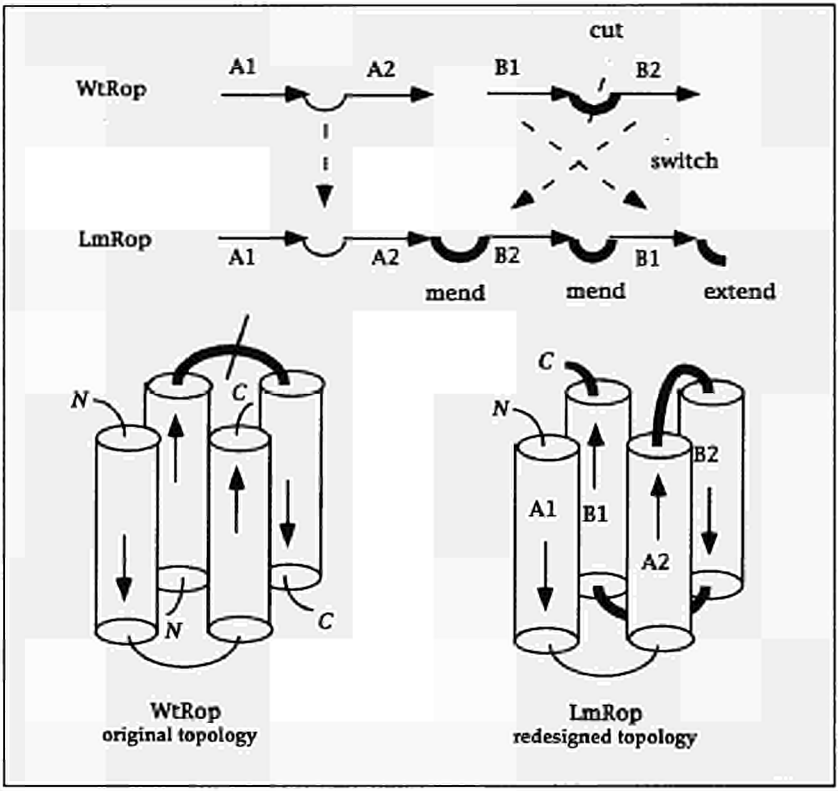


Figure 2: Reengineering topology of loop connections

Work on mutations in the core (CoreRop), building on the L41A and L48A mutants (Iraklion, Rome, Münster), has not yet led to crystal structures. The exhaustive exploration of core mutations was approached by an efficient Monte Carlo method. A third type of redesign aimed at functional properties: inserts into the engineered *Rop* monomer of functional loops imported from other proteins (LmRop-Achr/Gloop/Hghr). Finally, various metal binding sites have been engineered into the wild type *Rop* dimer. Work on these continues in 1994.

Screening and selection systems for mutants and epitope inserts (Roma)

The Roma group has given general support in molecular genetics to both the ferritin and *Rop* efforts. In addition, new genetic tools were developed to study the folding of four-helix bundles and helix-helix packing.

Analysis of packing constraints in the hydrophobic core of Rop

Following up on the work with cavity mutants, we have exploited a genetic test to ask whether the *Rop* core could accomodate larger residues. We constructed a large number of *Rop* mutants in which residues at positions 15(Ile), 19(Thr), 41(Leu) and 45(Ala) were substituted by a random combination of five amino acids with a hydrophobic side chain (Leu, Ile, Met, Phe, Val). We characterized

Table of selected engineered mutants of *Rop* protein (Braunschweig and Heidelberg)

protein(s)	*description*	*status*	*remark*
LmRop1/2 T21C/ K25C H44C	left-handed monomers	NMR, Xtals, Xray data	first successful non-cyclic reengineering of chain topology
LmRop2	designed for crystal heavy atom derivatives	purified on mg-scale	stable proteins, CD shows helical fold for T21C
RmRop1/2/ 3	right-handed monomers	expressed and purified, Xtal in one case	CD shows helical fold for RmRop2 and RmRop1-del-tail
CoreRop1-10	compensationg core mutants	several proteins expressed	designed to fill holes left in L41A and L48A
LmRop-Achr	insert of loop from acetylcholine receptor	purified on mg-scale	stable protein CD shows helical fold
LmRop-Gloop	insert of antigentic loop	expressed	unstable (proteolysis)
LmRop-Hghr	insert of loop from growth horm. receptor	purified on mg-scale	stable protein CD shows helical fold
MtRop1-12	designed metal binding sites	most variants expressed	prel. MALDI-MS evidence > metal binding

20 random mutants for their ability to dimerize at 37 and at 42°C. We find that most mutants no longer shows the characteristic pattern of small and large side chains. The mutants tested, compared to wild type, had an increase in the size of core side chains from 20% to about 60%. Surprisingly, most of the random side chain combinations are tolerated by the *Rop* structure. We conlcude that the *Rop* fold is mainly determined by the specific pattern of hydrophobic and hydrophilic side chains along its amino acid sequence, while precise and efficient packing of hydrophobic side chains in the core plays only a quantitative role in determining thermodynamic stability.

Linking an easily selectable phenotype to the folding of a common structural motif: Selection of rare turn mutations that prevent the folding of Rop

Similarly, we asked whether the *Rop* bend region contains important folding information. We extensively mutated three residues in the loop between the two α-helices. The characterization of random mutants indicated that the *Rop* native structure tolerates most amino acid types in these positions. To identify the rare

amino acid sequences that would prevent *Rop* from folding and/or dimerizing, we exploited the observation that *Rop* can functionally substitute the dimerization domain of the λ repressor. In fact, plasmids expressing a hybrid protein, formed by the amino terminal domain of the λ repressor covalently linked to *Rop*, confer immunity to λ infection on their hosts. This property depends on the ability of the *Rop* moiety to fold and dimerize. The analysis of approximately 1000 *Rop* mutants containing random amino acid sequences at position 30, 31 and 32 allowed us to identify 3 mutant *Rop* proteins that are not able to dimerize. In these mutants, the two helices are connected by the tripeptides VED, VPD and YPD in place of the wild type DAD (pos. 30, 31 and 32).

Filamentous bacteriophages as models for helix packing and vectors for peptide libraries

One of the aims of the project was to exploit α-helical bundles as scaffolds for the insertion of peptides and for the construction of peptide libraries. After achieving this with *Rop* and ferritin, we wanted to extend the scope of the technology by using protein vectors with the extra property of being able to bind their coding sequences. With this class of vectors, a selection procedure would enrich a specific property of the hybrid protein and at the same time would permit the cloning of its gene and the characterization of the modification that has caused the change of property. We chose the α-helical major coat protein of bacteriophage M13 (gene VIII product). We assembled a collection of 10^7 gene VIII, each of which typically synthesizes a different hybrid product formed by the native pVIII with an amino-terminal extension of nine amino acids. By probing this phage epitope library with a monoclonal antibody raised against a nonapeptide of human interleukin 1β, it was possible to select phages whose recombinant pVIII protein is recognized by the monoclonal antibody (Fig. 3). Their characterization revealed a consensus sequence (SND) that is present in all the selected peptides and is also found in the interleukin β1 peptide used to raise the antibody. This technology permits the

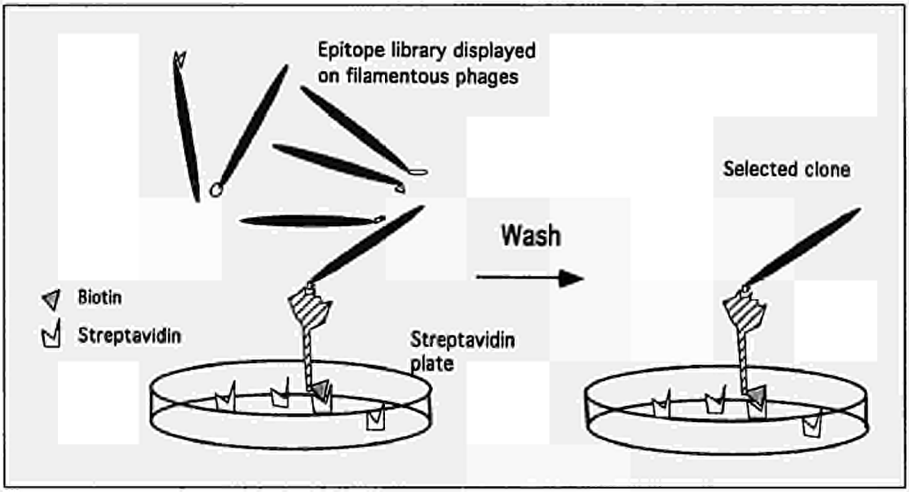


Figure 3: Selection system for epitope binding.
Selection from an epitope library of a filamentous phage that displays a peptide that binds a Fab fragment of immunoglobulin. The Fab fragment is linked to a plastic Petri dish via a biotin-streptavidin bridge.

mapping of immunogenic epitopes and the discovery of new ligands. Application in the diagnostic and pharmaceutical fields can be easily foreseen.

MAJOR SCIENTIFIC BREAKTHROUGHS

The folding, assembly and species variation of the iron-carrier ferritin was elucidated in great detail. A number of protein engineering techniques were shown to work in practice: first engineered crystal contact site, first non-cyclic alteration of chain topology, efficient genetic test of structural stability of a mutated dimer. Detailed predictive and design techniques for helical bundle proteins were developed. A wide variety of loop inserts led to several useful designed variants, including ferritin as a highly immunogenic carrier.

MAJOR COOPERATIVE LINKS

Milano, Rome and Sheffield collaborated closely on ferritin, including exchange of material and results. Rome, Iraklion, Heidelberg, and Braunschweig did likewise for *Rop* protein. Clones, purified protein and atomic coordinates were exchanged. All groups shared their results and had intensive discussion and planning meetings.

PUBLICATIONS

Joint publications

D.M. Lawson, P.J. Artymiuk, S.J. Yewdall, J.C. Livingstone, A. Treffry, A. Luzzago, S. Levi, P. Arosio, G. Cesareni, C.D. Thomas, W. Shaw and P.M. Harrison. Genetic engineering of crystal contacts leads to the structure solution of human H ferritin. Nature, **349**:541-544 (1991).

V.J. Wade, S. Levi, P. Arosio, A. Treffry, P.M. Harrison and S. Mann. Influence of site-directed modifications on the formation of iron cores in ferritin. J. Mol. Biol. **221**: 1443-1452 (1991).

S.C. Andrews, P. Arosio, W. Bottke, J-F. Briat, M. von Darl, P.M. Harrison, J-P. Laulhere, S. Levi, S. Lobreaux & S.J. Yewdall. Structure, Function and Evolution of Ferritins. J. Inorg. Biochem., **47**:161-174 (1992).

S. Levi, S.J. Yewdall, P.M. Harrison, P. Santambrogio, A. Cozzi, E. Rovida, A. Albertini and P. Arosio. Evidence that H- and L-chains have co-operative roles in the iron-uptake mechanism of human ferritin. Biochem. J., **288**: 591-596 (1992).

P. Santambrogio, S. Levi, P. Arosio, L. Palagi, G. Vecchio, D.M. Lawson, S.J. Yewdall, P.J. Artymiuk, P.M. Harrison, R. Jappelli and G. Cesareni. Evidence that a salt bridge in the Light chain contributes to the physical stability difference between Heavy & Light human ferritins. J. Biol. Chem., **267**: 14077-14083 (1992).

S. Mann, V. Wade, A. Treffry, S.J. Yewdall, P.M. Harrison, S. Levi and P. Arosio. Mapping of nucleation sites in ferritin: Influence of site-directed modifications on iron oxide mineralization. In: Chemistry and Biology of mineralized tissues, (Eds. H. Slavkin & P. Price), Excerpta Medica, Elseiver, Amsterdam, pp 399-408, (1992).

E.R. Bauminger, A. Treffry, A.J. Hudson, D. Hechel, N.W. Hodson, S.C. Andrews, S. Levi, I. Nowik, P. Arosio, J.R. Guest & P.M. Harrison. Iron incorporation into ferritins: evidence for transfer of monomeric Fe(III) between ferritin molecules and for the formation of an unusual mineral in the ferritin of *Escherichia coli*. Biochem. J. (in press)

P. Santambrigio, P. Pinto, S. Levi, A. Cozzi, E. Rovida, A. Albertini, P. Artymiuk, P.M. Harrison, and P. Arosio.On the mechanism of refolding and reassembly in human ferritin: effects of modifications near the 2-, 3- and 4-fold symmetry axes. (in preparation)

R. Jappelli, A. Luzzago, P. Tataseo, I. Pernice and G. Cesareni. Loop mutations can cause a substantial conformational change in the carboxy-terminus of the ferritin protein. J. Mol. Biol. 227:532-543 (1992).

G. Cesareni. Peptide display on filamentous phage capsids; a new powerful tool to study protein-ligand interaction. FEBS letters **307**: 66-70 (1992).

M. Kokkinidis, Vlassi, Y. Papanikolaou, M. Kotsifaki, A. Kingswell, D. Tsernoglou and H.-J. Hinz. Correlation between protein stability and crystal properties of designed *Rop* variants. Proteins **16**, 214-217 (1993)
C. Steif, P. Weber, J. Flossdorf, G. Cesareni, M. Kokkinidis, H.-J. Hinz. Subunit interactions provide a significant contribution to the stability of the dimeric 4-alpha-helical bundle protein Rop. Biochemistry **32**: 3867-3876 (1993)
S.C. Emery, M. Sagermann, W. Eberle, W. Klaus, H. Blöcker, P. Roesch, D. Tsernoglou & C. Sander. Reengineering protein topology. (Submitted).

Individual publications

M. Sagermann, M., Strukturanalyse an Funktions- und Topologievarianten des ColE1 *Rop* proteins (Structural analysis of functional and topological variants of the *Rop* protein), Diploma dissertation, University of Heidelberg (1990).
A. Pastore, P. Roesch, and C. Sander. The Structure of ColE1 *Rop* in Solution. W.Eberle, J. Biomol. NMR **1**, 71-82 (1991)
K. Hornischer. Herstellung und Untersuchung von potentiell metallbindenden Varianten eines 4-Helix-B ndel Proteins (Production and analysis of potentially metal binding variants of a 4-helix bundle protein). Diploma dissertation, University of Braunschweig (1991).
C.D. Paliakasis and M. Kokkinidis. Prot. The stability of the four-α-helix bundle motif. Engineering 4: 849-850 (1991).
C. Sander, M. Scharf and R. Schneider. Design of Protein Structures. In Protein Engineering, a Practical Approach, A.R.Rees, M.J.E. Sternberg and R. Wetzel eds., Oxford University Press, pp. 89-115 (1992).
H. Kresse. Herstellung und Untersuchung rekombinanter Rop-Protein-Varianten (Production and analysis of recombinant *Rop* protein variants). Ph.D. Disseration, University of Gøttingen (1993).
H.-J. Hinz, C. Steif, T. Vogl, R. Meyer, M. Renner, R. Ledermueller Pure and Appl. Chem. **63**: 947-952 (1993)
C. Sander. Design of protein strutures: helical bundles and beyond. Trends in Biotechnology **12**, 163-167 (1994)
Structural Design of Proteins. Sander, C., In: Concepts of Protein Engineering and Design, Eds. Wrede and Schneider, Walter De Gruyter, Berlin (1994).
CD. Paliakasis and M. Kokkinidis. Probability-based local alignment techniques. (submitted)
Restored heptad pattern continuity does not alter the folding of a 4-alpha-helical bundle. Vlassi et al., (submitted)

AREA B:

ENABLING TECHNOLOGIES

BIOTRANSFORMATION

New ways of biotransformation in non — aqueous systems for the synthesis of pharmaceuticals. Application of supercritical gases, organic solvents, liquid membranes and microemulsions (BIOT CT-900176)

COORDINATOR:
F.N. KOLISIS, National Hellenic Research Foundation, Athens, GR
PARTICIPANTS:
T. SCHEPER, Institut für Technische Chemie, Hannover Univ., Hannover, D
A. BALLESTEROS, Instituto de Catalysis, CSIC, Madrid, E
D. COMBES, Dép. Génie Biochimique et Alimentaire, INSA, Toulouse, F
U. MENGE, Gesellschaft für Biotechnologische Forschung, Braunschweig, D

BACKGROUND INFORMATION

Enzymes are favorite catalysts for the production of many fine-chemicals. However, the main obstacle for their extensive use in industrial processes is the problem of the low solubility of a large number of hydrophobic industrial important compounds in the water, the natural medium for enzyme action. In this project an extensive investigation has been undertaken in order to elucidate the engineering of the enzymes in various non-conventional media for the synthesis of hydrophobic products with great interest.

OBJECTIVES AND PRIMARY APPROACHES

1. Enzyme stability studies in supercritical carbon dioxide, effect of water content and effect of immobilization on a solid support.
2. Influence of $SCCO_2$ on enzyme activity and enantioselectivity.
3. Influence of reaction variables (e.g., pressure, medium composition, and temperature) on the enzymatic reactions.
4. A completed kinetic study of a model reaction (esterification of oleic acid with ethanol by an immobilized lipase) in a conventional organic solvent (n-hexane) compared to supercritical carbon dioxide.
5. Completed study of lipase specificity and kinetics in microemulsions.
6. Optimal reactor design.
7. Study of an integral reaction-separation process with a solvent recycled system.

RESULTS AND DISCUSSION

Various microbial lipases have been produced (Braunschweig lab) purified and characterized (Braunschweig, Hannover, Madrid labs). Also other esterases as subtilisine have been chemically modified and used in various synthetic reactions (Athens, Madrid labs).

A. Reactions in organic solvents

The chiral resolution of racemic mixtures of b-hyxdroxy esters were performed in organic phases with different lipases (Hannover lab). The reaction conditions have been optimized with 3-hydroxy octanoic acid methyl ester. Reactions were carrier out in batch and in continuous processes. In a solid bed reactor the continuous enantioselective separation of the substrate (99%eeS) was possible after short time. In the optimized system the resolutions of other 3-hydroxy esters were tested.

Aliphatic compounds reacted with lower enantioselectivity, only the substrates could be isolated in high enantiomeric purity.In contrast, aromatic 3-hydroxy esters were acylated by lipases with high stereoselectivity. The apparent kinetic parameters were determined and mechanistic studies were performed to obtain information about possible product/substrate inhibition.

Lipases were also used to investigate the regioselectivity of reactions concerning glycerides and acylated sucroses in organic solvents (Madrid lab). In this frame the transesterification between ethyl oleate and glycerol, or esterification between oleic acid and glycerine by stabilized lipase from *Candida antarctica* was studied. Transesterification was in all cases a better approach. When the reaction was carried out in heptane the presence of 3% water gave highest yields to mono and diolein; the formation of diolein can reach levels comparable to those of monolein. However, when the reaction was carried out in the presence of acetonitrile or acetone, and found to be much more selective to monolein. Similar results with smaller yields were obtained in the preparation of mono+distearin. Partially acylated sucroses prepared by various methods were used as substrates and regioselective acylation of xylopyranosides was accomplished.

B. Enzymatic reactions in microemulsions

The catalytic behavior of lipases from various microbial sources was investigated in anionic microemulsion systems stabilized by the surfactant molecule bis(2ethylhexyl)sulfosuccinate sodium salt (AOT) in a series of esterification reactions of various aliphatic alcohols with natural fatty acids (Athens lab).The effect of the nature of the fatty acids (chain length) and of the alcohols (primary, secondary or tertiary, chain length, cyclic structures) on lipases from *Rhizopus delemar, Rhizopus arrhizus, Pseudomonas cepacia* and *Penicillium simplicissimum* activities was investigated, in relation to the reverse micellar structure. The lipases tested showed a selectivity regarding the structure of the substrates used. *P. simplicissimum* lipase showed higher reaction rates in the esterification of long chain alcohols as well as secondary alcohols. Primary alcohols had a low reaction rate and tertiary a very slow rate of esterification. Long chain fatty acids were better catalyzed as compared to the shorter ones. *R. delemar* and *R. arrhizus* lipases showed a preference for the esterification of short chain primary alcohols, while the secondary alcohols had a low rate of esterification and the tertiary ones could not be converted. The reaction of medium chain length fatty acids was also better catalyzed than in the case of the long ones. *P. cepacia* shows a preference for palmitic and caprylic acid as well as for propanol, while reactions involving cyclic alcohols can not be catalyzed at all. The observed lipase selectivity appeared to be related to the localization of the enzyme molecule within the micellar microstructure, due to the hydrophobic/hydrophilic character of the protein. The reverse micellar structural characteristics as well as the localization of the enzyme, were examined by fluorescence quenching measurements and spectroscopical studies. Kinetic studies of the esterification of octanol by lauric acid catalyzed by *P. cepacia* lipase showed that the reaction follows a ping-pong bi bi mechanism with inhibition by both substrates. The apparent kinetic parameters of the reaction were found to be $Km_{octanol}$ = 310 mM, $Km_{laurate}$ = 78 mM and a Vmax = 250 μmol $min^{-1}mg^{-1}$. The same microemulsion system was used for the synthesis of mono and diglycerides from glycerol and lauric acid, which was successful at very low w values. The catalytic behavior of *P. cepacia* lipase was also studied in esterification reactions performed in a non-ionic microemulsion system formulated by the non-ionic surfactant Tetraethyleneglycoldodecylether ($C_{12}E_4$). The optimum activity was found at about w_o=8. The apparent values of $Vmax^{app}$ and Km^{app} for octanol

were calculated and found to be 100 $\mu molmin^{-1}mg^{-1}$ and 76 mM respectively. In addition it was found that the lipase from *Penicillium simplicissimum* catalyzes the stereospecific esterification of menthol with fatty acids. Microemulsions appear to be an effective and fast system for racemic resolution of alcohols. Finally the important parameters of the reactions, such as temperature, pH and water content (w_o) were identified for all the above mentioned lipases.

C. Enzymic reactions in supercritical carbon dioxide

1. Toulouse lab.

In this field the following investigations were carried out.

i) The lipozyme is stable in $SCCO_2$ in a range of pressure from 13 to 18 MPa. Stability is similar in $SCCO_2$ and in n-hexane and the results obtained over a period of 6 days lead to the possibility of using $SCCO_2$ in a continuous enzyme reactor. Water has been demonstrated as an important denaturing factor.

ii) The first completed study of kinetic behavior of enzymatic reactions in $SCCO_2$ compared to a more conventional medium (n-hexane) was carried out. The kinetic constants (Michaelis and inhibition constants) and the mechanism of action of Lipozyme (ping-pong BiBi with inhibition by one substrate, ethanol) has been determined. The same behavior is observed in both media.

iii) A continuous reaction separation process has been developed and good results have been obtained with one fixed bed reactor coupled with four separators.

iv) The limitation of the solubilization of polar substrates in organic solvents and $SCCO_2$ has been overcame with the use of complexation with phenylboronic acid (PBAC) and immobilization of polar substrate on silica gel. In both cases, the feasibility has been demonstrated through the esterification of oleic acid with glycerol or D-fructose. Significant conversion was observed in both cases (more than 50% for glycerol and more than 30% for fructose).

2. Hannover lab.

A complete reactor system was set up in which an on-line monitoring of the reactions was possible. The results show, that reactions such as the chiral resolution of racemic 3-hydroxy esters can be performed in $SCCO_2$ with similar or even better results as in comparable organic phases (e.g. hexane or heptane). These research points were studied in detail and important information about the reactor (especially for on-line monitoring), enzyme stabilization, process conditions and organization, and about the reaction medium engineering were gained.
Enzymes immobilized on polymer supports via epoxide-reactions and enzymes stabilized via affinity reactions with lectines were used during the studies. In all studies with non-glycosylated enzymes, lipases from *Pseudomonas cepacia* were used. The most important results for these studies can be summarized as:

i) vinylacetate inhibits the esterification reactions with increasing concentrations but also increases (in contrary to organic solvents) the enantioselectivity of the reaction;

ii) the water concentration is of immense importance for the whole reaction; above a certain critical concentration (differs from system to system) the reaction rate as well as the enantioselectivity is decreasing, thus a reduction of water by addition of molecular sieves is optimal;

iii) the density of the reaction medium (influenced by pressures and reaction temperature) is of importance, since the solubility of water is a function of it;

iv) if the density is constant, changes in temperature or pressure do not affect the enzyme activity and enantioselectivity in a certain range;

v) during medium engineering a positive effect (on reactivity and selectivity) were found for unpolar additives, while a negative effect was found for polar additives,

vi) in direct comparison with hexane, a better enantioselectivity was found in the organic phase, while a higher yield was found for the $SCCO_2$;

vii) enzyme activity decreases during reaction in $SCCO_2$, and also during the pressurizing and depressurizing steps of the whole process.

3. Braunschweig lab.

The stability of lipases from *Rhizopus miehi, Rhizopus oryzae, Penicillium camembertii*, and *Pseudomonas cepacia* in $SCCO_2$ has been determined using a multi-reactor system. The loss of activity can be attributed to high losses of water solubilized in the gaseous CO_2 phase. Immobilization of the enzyme gives a significant protection as compared to purified lipase. The addition of small amounts of water or polar organic solvents destabilized lipases immobilized on hydrophobic supports more than those immobilized on hydrophilic ones. In the presence of polar cosolvents with a dielectric constant of $\varepsilon < 2$ *R. miehi* and *P. camembertii* lipases were completely inactivated during incubation in $SCCO_2$ if bound on duolite. In contrast *P. cepacia* lipase was affected only to a minor extent by the immobilization support. After a proper selection of lipase source and cosolvent the $SCCO_2$ bioreactors can be run more than a week without loss of enzyme activity.

MAJOR SCIENTIFIC BREAKTHROUGHS

1. Isolation, purification, and crystallization of new lipases.
2. Exact knowledge about medium engineering.
3. On-line analysis of enzymatic reactions in $SCCO_2$.
4. Optimization of model esterification reactions in all non-conventional systems used.
5. Development of novel chiral GC-analysis for all synthons.
6. Development of new monitoring devices in non-conventional media (optodes).
7. Development of novel enzyme stabilization methods (patent pending).
8. The reaction-separation coupling in $SCCO_2$ with solvent recycling.
9. The solubility limitation of polar substrates in $SCCO_2$ is overcame.
10. Microemulsions is a fast system for racemic resolution of alcohols.

MAJOR COOPERATIVE LINKS

Participation in the Comett-programm (92/1/6903/Ca) on the topic of the BRIDGE project. Five workshops with a total of 16 sessions were carried out in Greece, France and Germany. A patent will be submitted based on the results of the BRIDGE project within the Value II programme, and a Value II proposal was submitted (AC-162/93-D) for a possible commercialization of the results. **Industrial support:** Merck Darmstadt supported the project by supplying fine chemical to all partners (total value in 1993: 5000ECU). **Exchange of materials and staff:** Enzyme preparations and different immobilizats were exchanged among the partners. Also various common experiments were performed and visits and joint meeting of the partners took place.

PUBLICATIONS

Joint publications

T. Scheper, U. Bornscheuer, A. Capewell, A. Herar, H.G. Hundeck, H.H. Meyer, F.N. Kolisis and F. Schubert (1992). Studien zu enzymatischen reaktionen in wassrigen, organischen und uberrkritischen phasen. DECHEMA — Karlsruhe,1-3 June.

A. Capewell, U. Bornscheuer, T. Scheper, H.H. Meyer and F.N. Kolisis (1992). A comparison of enzymatic reactions in aqueous, organic and supercritical phases. *Ann. N.Y. Acad. Sci.* **672**, 336-342.

U. Bornscheuer, A. Herar, L. Kreye, V. Wendel, A. Capewell, H.H. Meyer, T. Scheper and F.N. Kolisis (1993). Factors affecting the lipase catalyzed transesterification reactions of 3-Hydroxyesters in organic solvents. *Tetrahedron Asymmetry.*, **4** (5), 1007-1016.

H. Stamatis, A. Xenakis, S. Sztajer, U. Megne and F.N. Kolisis (1992). Studies on the specificity of *penicillium simplicissimum* lipase catalyzed esterification reactions in microemulsions. *Elsevier Science Publishers., Ed. J. Tramper.* Progress in Biotechnology vol. **8**, 733-738.

T. Scheper, U. Bornscheuer, A. Capewell, A. Herar, H. G. Hundeck,H. Meyer, F. Schubert and F.N. Kolisis (1992). A comparison of enzymatic reactions in aqueous, organic and multiphase systems. *Elsevier Science Publishers., Ed. J. Tramper.* Progress in Biotechnology vol. **8**, 37-44.

Koening K., Menge U. and Ballesteros A. (1992) Enzyme activity, stability and stabilization in organic solvents and supercritical carbon dioxide. Symposium no stability & stabilization of enzymes, Maastricht, Nov. 22-25.

H. Stamatis, A. Xenakis, U. Menge and F.N. Kolisis (1993). Kinetic study of lipase catalyzed esterification reactions in w/o microemulsions. *Biotechnology and Bioengineering*, **42**, 931-937.

H. Stamatis, A. Xenakis, U. Bornsheuer, T. Scheper, U. Menge and F.N. Kolisis (1993). *Pseudomonas cepacia* lipase: esterification reactions in AOT microemulsion systems. *Biotechnol. Lett.*, **15**, 703-708.

A. Ballesteros, U. Bornscheuer, A. Capewell, D. Combes, J-F. Condoret, K. Koening, **F.N. Kolisis**, A. Marty, U. Menge, T. Scheper and A. Xenakis (1993) Enzymes in non-conventional phases. *Biocatalysis*, (in press).

U. Bornscheuer, O-W. Reif, R. Lausch, R. Freitag, T. Scheper, U. Menge and F.N. Kolisis (1993). Purification, Crystallization and Characterization of lipase from *Pseudomonas cepacia. Biochem. Biophys. Acta.*, (in press)

Individual publications

A. Xenakis, T. Valis and F.N. Kolisis (1991). Microemulsions as a tool for enzymatic studies.The case of Lipase. *Progr. Colloid Polym. Sci.*, **84**, 508 — 511.

T. Valis, A. Xenakis and F.N. Kolisis (1992). Comparative studies of Lipase from *Rhizopus delemar* in various microemulsion systems, *Biocatalysis*, **6**, 267-276

G. Kodelia and F.N. Kolisis (1992). Studies on the reaction catalysed by subtilisin for the acylation of flavonoids with fatty acids in organic solvents. *Ann. N.Y. Acad. Sci.*, **672**, 451-457.

H. Stamatis, A. Xenakis, M. Provelegiou and F. N. Kolisis (1993). Comparative studies on the selectivity of lipases in microemulsions for esterification reactions. The role of enzyme localization. *Biotechnology and Bioengineering*, **42**, 103-110

A. Xenakis, H. Stamatis, A. Maliaris and F.N. Kolisis (1993). Effects of alcohols on the structure of AOT reverse micelles with respect to different enzyme activity. Progr. Colloid Polym. Sci., **93**, 373-376.

H. Stamatis, A. Xenakis and F.N. Kolisis (1993). Enantiomeric specificity of a lipase from *Penicilium simpicissimum* in the esterification of menthol in microemulsions. *Biotechnol. Lett.*, **15**, 471-476.

H. Stamatis, A. Xenakis, F.N. Kolisis and A. Malliaris (1994) Lipase localization in water-in-oil microemulsions studied by fluoresence energy transfer. *Progr. Colloid Polym. Sci.*, (in press).

G. Kodelia, K. Athanasiou and F.N. Kolisis (1994). Cytogenetic effects of rutin and enzymatically synthesized butyryl-rutin ester on mammalian cells in culture. *Appl. Biochemistry and Biotechnology*, **44** (2), 000-000.

H. Stamatis, A. Xenakis and F.N. Kolisis (1994). A novel approach for lipase reuse and product recovery in esterification reactions performed in microemulsions. *Ann. N.Y. Acad Sci.* (in press).

C. Tsitsimpikou, D. Voutou, P. Christakopoulos, D. Kekos, B.J. Macris and F.N. Kolisis Studies of the effect of organic solvents on the stability of β-glucosidase from fusarium oxysporum. *Biotech. Lett.*, **16**, 57-62.

H. Stamatis, A. Xenakis, S Dimitriadis and F.N. Kolisis (1994) Catalytic behavior of *Pseudomonas cepacia* lipase in w/o microemulsions. *Biotechnol Bioengineer.*, (in press).
U. Bornscheuer, S. Schapohler, T. Scheper and K. Schugerl (1991). Influences of reaction conditions on the enantioselective transesterification using *Pseudomonas cepacia lipase. Tetrahedron Asymmetry* 2(10), 1011-4.
U. Bornscheuer,S. Schapohler, T. Scheper, W.A. Konig and K. Schugerl. Application of enantioselective capillary gas chromatography in lipase catalyzed transesterification reactions in organic media. *J. Chromatog.*, (in press).
M.A Cruces, C. Otero, M. Bernabe, M. Martin-Lomas and A. Ballesteros (1992). Enzymatic preparation of acylated sucroses. *Annals N.Y. Acad. Sci.*, **672**, 436-443.
M. L. Rua, T. Diaz-Maurino, C. Otero and A. Ballesteros (1992). Isoenzymes of lipase from *Candida cylindracea*. Structural studies on the carbohydrate chains. *An. N.Y. Acad. Sci.*, **672**, 19-23.
E. Pastor, C. Otero and A. Ballesteros (1992). Hidrolisis de grasas catalizada por lipasa immobilizada sobre soportes siliceos. Proc. XIII Simposio Iberoamericano de Catalisis, S. Mendioroso (ed), 581-584.
F.J. Plou, C. Otero and A. Ballesteros (1992). Preparacion de mejores biocatalizadores por modificacion quimica. Proc. XIII Simposio Iberoamericano de Catalisis, S.Mendioroz ed), 577-580.
Lopez R. and A. Fernadez-Mayorales (1992) Controlling yield and regioselectivity in the enzymatic synthesis of β-D-galactopyranosyl-β-D-xylopyranosides. *Tetrah. Lett.*, **33**, 5449-5452.
M.L. Rua, T. Diaz-Maurino, V.M. Fernadez, C. Otero and A. Ballesteros (1993) Purification and characterization of two distinct lipases from *C. cylidracea. Bioch. Bioph. Acta*, **1156**, 181-189.
R. Lopez, C.Perez, A. Fernadez-Mayorales, S. Conde (1993) Enzymatic transesterification of alkyl 2,3,4-tri-O-acyl-β- D- xylopyranosides. J. Carbohyd. Chem., **12**, 165-172.
F.J. Plou and A. Ballesteros (1994) Acylation of substilisine with long fatty acyl residues affects its activity and thermostability in aqueous medium. *FEBS Lett.*, **9**, 200-204.
M.L. Roua and A. Ballesteros (1994) Rapid purification of two lipase isoenzymes from *Candida rugosa, Biotechnol. Techniq.*, **8**, 21-26.
E. Pastor, A. Ballesteros, F.J. Plou and C. Otero (1994) Enzymatic synthesis of oleins in organic media. *Annals N.Y. Acad. Sci.* (in press).
F.J. Plou, M.A. Cruces, M. Bernabe, M. Martin-Lomas, J.L. Parra and A. Ballesteros (1994) Enzymatic synthesis of partially acylated sucroses. *Annals N.Y. Acad Sci*, in press.
Marty A., Moreau V., Combes D. and Condoret J.S. (1991) Production en continu d'ester d'acides gras par biocatalyse dans d'acides gras par biocatalyse dans du dioxyde de carbone supercritique. *Les Fluides Supercritique: Caractéristiques et Applications,* M. Perrut Ed., INP Lorraine, 287-292.
J.S. Condoret, A. Marty and D. Combes (1991). Etude cinetique de catalyse enzymatique en fluide supercritique. Colloque Bioconversion et Synthese Organique, Grasse. France, 22-24 Mai.
D. Combes, W. Chulalaksananukul, A. Marty and J.S. Condoret (1991). Comparison of immobilized lipases behavior in supercritical CO_2 and n-hexane. RSC conference: Lipases: structure function and application in biotransformations. University of Warwick, Coventry, U.K., July 16-18.
A. Marty, V. Moreau, D. Combes and J.S Condoret (1991). Production en continu d'esters d'acide gras par biocatalyse dans du dioxyde de carbone supercritique. 2^e Colloque sur les Fluides Supercritiques, Paris, France, 16-17 October.
A. Marty, W. Chulalaksananukul, R.M. Willemot and J.S. Condoret (1992). Kinetics of the lipase-catalyzed esterification in supercritical CO_2. *Biotechnol. Bioeng.*, **39**, 273-279.
Marty A., Combes D. and Condoret J.S. (1992) Fatty acid esterification in supercritical carbon dioxide. *Elsevier Science Publishers., Ed. J. Tramper.* Progress in Biotechnology vol. **8**, 425-432.
W. Chulalaksananukul, J.S. Condoret and D. Combes (1993). Kinetics of geranyl acetate synthesis by lipase catalyzed transesterification in n-hexane. *Enzyme Microb. Tech.*, **15**, 691-698.
Castillo E., Marty A., Combes D. and Condoret J.S. (1994) Polar substrates for enzymatic reactions in supercritical CO_2: How to overcome the solubility limitation. *Biotechn. Lett.*, **16**, 169-174.
Marty A., Combes D. and Condoret J.S. (1994)Continuous reaction-separation process for enzymatic esterification in supercritical carbon dioxide. *Biotechn. Bioeng.*, in press.

Patents

Rua M.L. Rodrigez and A. Ballesteros (1992),Procedure for the purification of two isoenzymes lipase from *Candida rugosa*. Spanish Patent No 9201378 (intern. ext. No PCT/ES93/00058).

Characterization and surface properties of semi-synthetic redox enzymes for their applications in biosensor devices (BIOT CT-910279)

COORDINATOR:
A.F. BÜCKMANN, Gesellschaft f. Biotechnologische Forschung mbH, Braunschweig, D
PARTICIPANTS:
J.V. BANNISTER, Cranfield Institute of Technology, Bedfordshire, GB
C. van DIJK, Agrotechnological Research Institute, Wageningen, NL
S.G. MAYHEW, University College Dublin, Dublin, IRL
D. COCCO, University of Cagliari, Cagliari, I
J. PRENOSIL, ETH Zürich, Zürich, CH
G. GOTTSCHALK, Wiss-Trans Umwelt, Göttingen, D

BACKGROUND INFORMATION

Semi-synthetic redox enzymes, e.g. apo-oxidases with covalently bound FMN or FAD analogues and redox mediators or dehydrogenases with covalently bound NAD(H), that catalyse for biosensor development interesting reactions, may show favourable properties under the stress conditions when integrated in amperometric or electro-optic biosensor systems.

OBJECTIVES AND PRIMARY APPROACHES

Chemically modifying redox enzymes of bioanalytical interest (oxidases and dehydrogenases) by covalent attachment of coenzymes (FMN, FAD, NAD(H)) and redox mediators to their apo-enzymes and testing the properties of such adducts in connection with their continuous performance in biosensor systems.
Screening of thermophile bacteria exhibiting activities of thermostablle oxidases.

RESULTS AND DISCUSSION

A. FAD analogues

The aim is the development of chemistry to introduce in the adenine moiety of FAD functional groups (-NH_2 , -COOH or -N_3) at variable distance to the C(6)- or C(8)-position determined by the length of the spacer inbetween. Final aim is to couple such FAD analogues covalently to the apo-enzyme form of FAD dependent oxidases of interest for biosensor development in such a way that the enzymatic activity is maintained. Beside trial and error strategy in case of oxidases without known structure (e.g. for D-aminoacid oxidase and NADH oxidase from *Thermus aquaticus*) structural data will be used for the design of suitable FAD analogues with respect to spacer length for oxidases with known three dimensional structure (e.g. for glucose oxidase from Aspergillus niger and *Penicillium amagasakiensis* and NADH oxidase from *Thermus thermophilus*).

1. Synthesis of functionalized FAD from FMN (riboflavin-5'-phosphate) and functionalized AMP and preparation of an enzymatically active covalent FAD-D-aminoacid-oxidase adduct

The preparation of pure starting compound FMN from a commercial product containing riboflavin-4'-phosphate (6%) and riboflavin-3'-phosphate (10%) and

riboflavin (10%) in quantities of 100-200 mg by fast FPLC procedure is now routine.

A major chemical problem encountered in the first year of the project was the coupling of FMN to N^6-(carboxyalkyl)-AMP by diphosphate coupling after activation of the 5'-phosphate moiety of the latter compound. It turned out that the free carboxyl group interfered severely with respect to this coupling, presumably, due to an own strongly competitive simultaneous activation, since the coupling of FMN and AMP to FAD under equal conditions gave no problems.

A routine procedure has been developed for the the synthesis of N^6-(carboxyalkyl)-FAD by diphosphate coupling of FMN and the heptylester of N^6-(carboxyalkyl)-AMP to the heptylester of N^6-(carboxyalkyl)-FAD and subsequent enzymatic deprotection by lipase (e.g. synthesis of 200 mg N^6-(6-carboxyhexyl)-FAD with 30% overall yield).

For the synthesis of N^6-(aminohexyl)-FAD a similar procedure is currently under investigation starting out by diphosphate coupling of FMN and N^6-(N-phenylacetylhexyl)-AMP to N^6-(N-phenylacetylhexyl)-FAD and subsequent enzymatic deprotection to N^6-(aminohexyl)-FAD by immobilized penicillin-G-acylase as exemplified for N^6-(6-aminohexyl)-FAD. With respect to the latter synthesis, N-phenylacetylalkylamine had to be synthesized first and subsequently reacted with 6-chloropurine-riboside-5'-phosphate to N^6-(N-phenylacetylalkyl)-AMP. Unexpectedly, it turned out that the synthesis of N-phenylacetylalkylamine from phenylacetylchloride and the related diaminoalkan (1 mol/1 mol) was complicated. By ^{1}H-NMR it could be demonstrated that the free NH_2 group of N-phenylacetylalkylamine reacts preferably with phenylacetylchloride even in the presence of an excess of diaminoalkane, since the major product was always N,N'-diphenylacetylalkane as was investigated for the reaction of phenylacetylchloride and 1,6,-diaminohexane in THF.

An alternative method has been developed for the synthesis of N-phenylacetylalkylamine by preparing first N-BOC,N'-phenylacetylalkane by reaction of phenylacetylchloride and N-BOC-diaminoalkane and subsequent removal of the protecting BOC group by reaction with trifluoric aceticacid (e.g. 1.4 g N-phenylacetylhexylamine has been prepared with approx. 20% overall yield from phenylacetylchloride and N-BOC-1,6-diaminohexane).

For the first time, an enzymatically active covalent FAD-oxidase adduct has been synthesized.
N6-(6-carboxyhexyl)-FAD, activated by O-(N-succinimidyl)-N,N,N',N'-tetramethyluroniumtetrafluoroborate has been covalently attached to apo-D-aminoacid oxidase from pig kidney. Based on the results of dilution experiments and extrapolation, a procedure has been developed to obtain an adduct with 1 molecule FAD analogue/subunit. By comparing the molecular weight of apo-D-aminoacid oxidase and this FAD-D-aminoacid oxidase adduct by the electro-spray MS-technique this could be verified. The difference in molecular weight, 893, the molecular weight of the coupled FAD analogue, is illustrated in Fig. 2. The adduct did not loose the FAD analogue under the harsh conditions for preparing the apo-enzyme. Since the activity of this adduct was comparable to that of the native enzyme, it can be concluded, that the covalently bound FAD analogue molecule must be structurally positioned in the subunit similarly to native FAD in the holo-enzyme.

Currently, conditions are investigated for the crystallization of this FAD analogue-D-aminoacid oxidase adduct. Perhaps this adduct is better crystallizable than native

D-aminoacid oxidase for which the protein structure could not be determined up to now.

The reaction pathways of the new synthesis of N^6-(carboxyalkyl)-FAD (chemical modification in combination with enzymatic deprotection of the functional group) and the synthesis of the covalent FAD-D-amino-acid adduct have been depicted in Fig. 1.

2. Functionalization by chemical modification of intact FAD

N^6-(2-aminoethyl)-FAD, starting compound for further functionalization (e.g. introduction of a functionalized PEG spacer) can be synthesized by an in 1991 patented method: alkylation of FAD with ethyleneimine to introduce a 2-aminoethyl group at the N(1)-position of the adenine moiety followed by unexpectedly fast Dimroth rearrangement to N^6-(2-aminoethyl)-FAD under mild aqueous conditions (pH 6.0—6.5, 50°C, reaction time approx. 6 h, for the reaction pathway of this new FAD modification see BRIDGE Progress Report 1993 p. 81-92). By a combination of anion exchange chromatography and reversed phase HPLC 120 mg pure N^6-(2-aminoethyl)-FAD have been obtained with an overall yield of 22%.

N^6-(2-aminoethyl)-FAD has been further modified by reaction with N-hydroxysulfosuccinimidyl 4-azidobenzoate and sulfosuccinimidyl 4-(p-azidophenyl)butyrate to introduce an aromatic azide group at the 2-aminoethyl

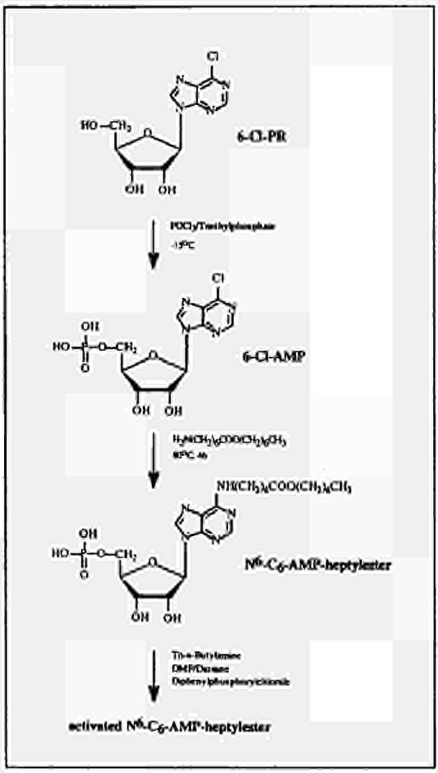


Figure 1 a

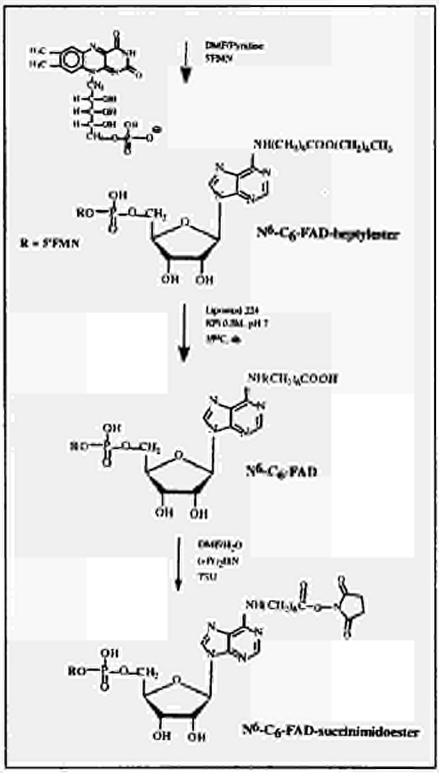


Figure 1 b

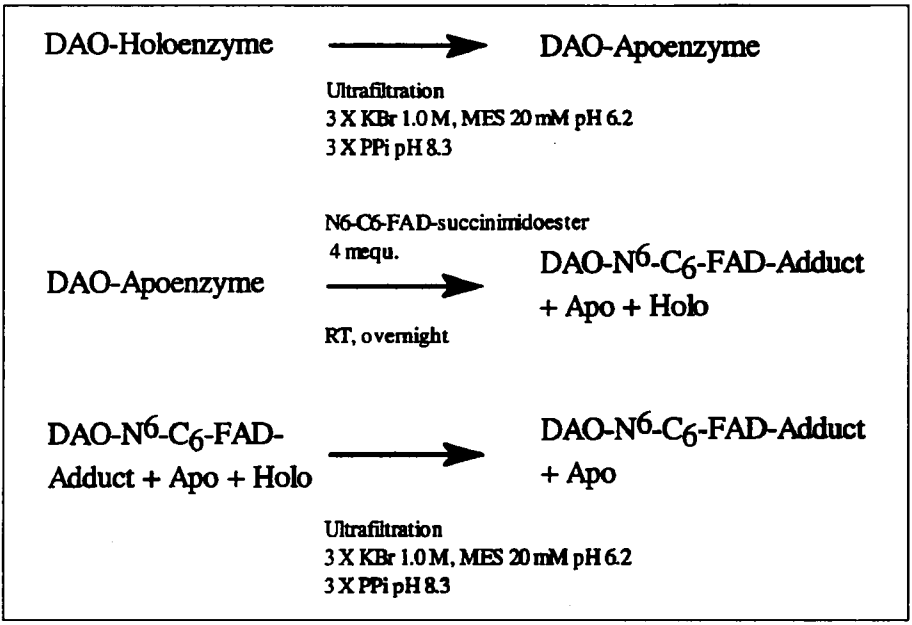


Figure 1 c

Fig. 1. Synthesis of N^6-(6-carboxyhexyl)-FAD (N6-C6-FAD) and preparation of the covalent N^6-(6-carboxyhexyl)-FAD-D-aminoacid oxidase (pig kidney) adduct (N6-C6-AMP-heptylester: heptylester of N^6-(6-carboxyhexyl)-AMP; N^6-C6-FAD-heptylester: heptylester of N^6-(6-carboxyhexyl)-FAD; Lipomod 224: lipase used for enzymatic deprotection)

group by amide bonding. These azido-FAD derivatives will be used to investigate the possibility of attaching FAD analogues to apo-oxidases by the special chemistry used in photoaffinity labeling.

Unfortunately, the enzymic synthesis of 8-Br-FAD from 8-Br-ATP and FMN, as reported in the last report, was found to be in error. Purified FAD-synthetase from *Brevibacterium ammoniagenes* is not active with 8-Br-ATP. On the contrary, 8-azidoadenine-FAD could be synthesized according to this enzymatic strategy with the starting compounds FMN and 8-azido-ATP

The possibility of synthesizing FAD that is brominated at the 8-position of the adenine moiety by chemical bromination of FAD was investigated. Reaction of FAD with Br_2 (up to 2x excess) for 12 h at 20°C led to the formation of a product distinguishable from FAD by TLC. The product was purified by chromatography on DEAE-cellulose and on Sep-Pak C_{18}. The product has light absorption and fluorescence properties slightly different from those of FAD. Treatment with FAD hydrolase (*Naja Naja* venom) generates a flavin identical to FMN (binding by an FMN-specific apoflavodoxin) and a modified form of AMP that has spectroscopic properties similar to those of 8-Br-AMP (absorption maximum at 265 nm).

The fluorescence emission of the modified FAD is not affected by the apo-enzyme of D-amino acid oxidase under conditions in which the fluoresence of FAD is quenched. This suggests that either the flavin is not bound by the apoenzyme or that it is bound only very weakly.

B. FMN and Isoalloxazine analogues

1. Functionalization of FMN

Currently, the modification of FMN is being investigated by trying to react the 5'-phosphate moiety with the -OH or -NH_2 group (P-O- or P-N-bond formation) of linear compounds with functional groups (-COOH or -NH_2 in the protected form).

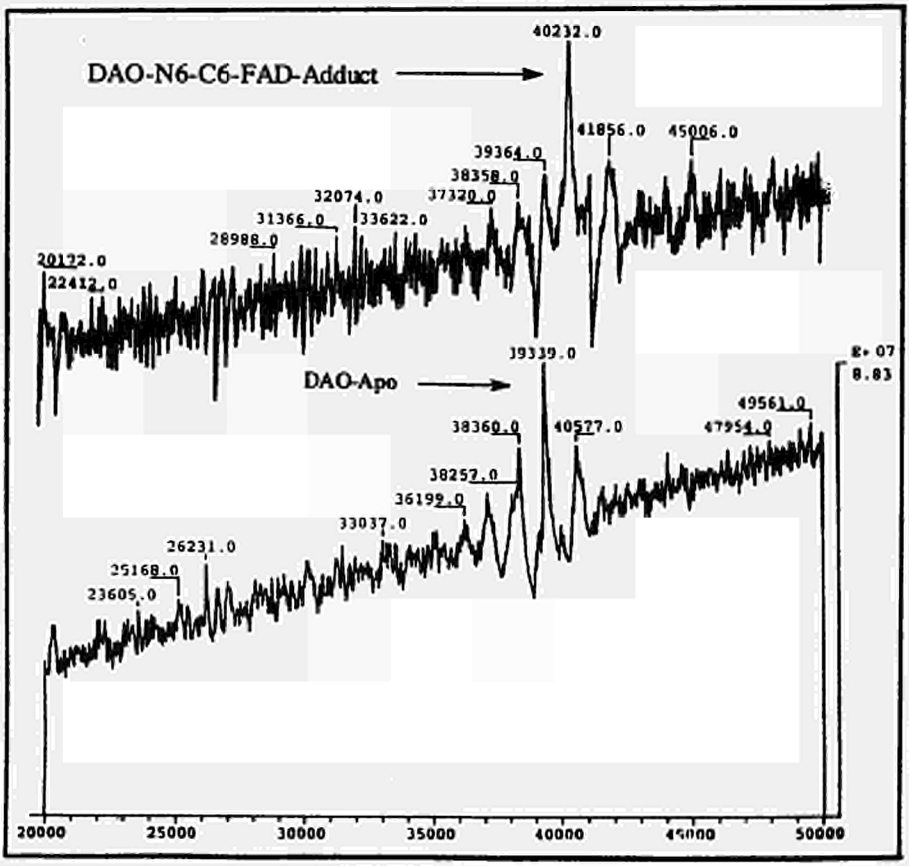


Figure 2. Electro-spray MS spectra of the N^6-(6-carboxyhexyl)-FAD-D-aminoacid oxidase adduct and apo-D-aminoacid oxidase from pig kidney.

2. Functionalization of isoalloxazine

Recently, 7,8-dimethyl-10-formylmethyl-isoalloxazine in the acetal form was synthesized by oxidation of riboflavine with periodic acid (4.6 g with an overall yield of approx. 60%).

Via the aldehyde group at the N(10)-position carboxylation have been carried out with linear compounds of the type NH_2-(CH_2)n-COOH (n = 1-7) by Schiff base formation, subsequent reduction with $NaCNBH_3$ and purification by FPLC. Such simple flavin analogues may be useful starting compounds for the synthesis of

active covalent flavin-oxidase adducts or may replace e.g. ferrocene analogues as redox mediators, for example, covalently coupled to oxidases.

C. Redox studies concerning electron acceptors interacting with oxidases

Oxidase electrodes have become a useful analytical tool, because of the unique combination of enzyme selectivity and the ease of automation using electrochemical methods. For continuous measurements, the coenzyme FMN or FAD of the oxidases (immobilized within the sensor) have to be regenerated in the oxidized form after each catalytic reaction step. Oxygen is the natural mediator to oxidize the in the reaction reduced FMN or FAD leading to H_2O_2 formation, that can be amperometrically detected. The classical enzyme electrodes based on this measurement principle rely on the presence of unlimited oxygen supply. Unfortunately, the regeneration of oxidases cannot effectivily be accomplished by direct oxidation of the redox centre of FMN or FAD at the electrode surface because the protein shell protects this redox centre and suppresses easy electron tunnelling. The discovery of the use of artificial electron transfer mediators to regenerate oxidases has stimulated the development of a new class of oxygen independent biosensors that may also operate under anaerobic conditions.

Prior to electron transport studies with respect to artificial electron transfer mediators and covalent adducts of target apo-oxidases and analogues of FMN and FAD, such studies have been carried out with the unmodified form of glucose oxidase and two NADH oxidases.

I. Glucose oxidase from Aspergillus niger

Consolidated results will be given for the reaction and interaction of artificial electron acceptors with quinoid structures and bipyridyl moieties with respect to this common glucose oxidase known in the field of enzyme technology and clinical and analytical chemistry.

Table 1.
Effect of pH on the pseudo first order rate constant and on the K_m value of PQQ

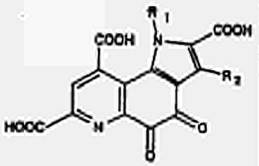
pH	$k_{cat}(sec^{-1})$	K_m (μM)
2.6	2.5	19.2 ± 1.6
2.8	5.4	37.2 ± 1.2
3.0	7.4	96.2 ± 1.8
3.2	5.6	185.7 ± 2.2
3.5	3.4	506.2 ± 1.6

a. Pyrrolo-quinoline quinone (PQQ) as electron acceptor.

PQQ was investigated as a possible electron acceptor for glucose oxidase over a wide pH range (2.4 — 8.0). The products of the reaction, pyrrolo-quinoline quinol and gluconic acid were identified and quantified by absorption spectroscopy and HPLC, and the stoichiometry of the reaction was determined. Kinetic studies on

the reaction at different pH values show that PQQ acts as oxidant for glucose oxidase only at acidic pH values with a maximum rate at pH 3.0. When the pH decreases from 5.0 to 3.0 an increase of the reaction rate with approximately three orders of magnitude was observed. The turnover number K_{cat} and K_m were determined for the pH range 2.6 — 3.5 (Table 1).

Table 2. Biomolecular rate constants for the reaction of PQQ derivatives with glucose oxidase

	$k^{a)}$, $(M^{-1}.s^{-1})$	%
R_1 = -H; R_2 = -H	$(1.45 \pm 0.06)x10^5$	100.0
R_1 = -H; R_2 = $-CH_3$	$(9.31 \pm 0.07)x10^4$	64.2
R_1 = $-CH_3$; R_2 = -H	$(2.58 \pm 0.10)x10^3$	1.8
R_1 = $-C_2H_5$; R_2 = -H	$(1.86 \pm 0.08)x10^3$	1.2

a) The rate constants were determined at pH 3.0, the optimum pH for the reaction of PQQ with glucose and glucose oxidase.

For pH values higher than 3.5, linear variations of the initial reaction rate with the PQQ concentration were obtained. This indicates that above pH 3.5 the rate determining step of the reaction is probably the quinone-enzyme complex formation. The decrease of K_m upon decreasing the pH suggests a stronger association of glucose oxidase with the protonated form of PQQ of which the most acidic carboxyl group has a pK_a of 2.0. It is generally accepted that the redox properties of PQQ are induced by its o-quinone moiety. If PQQ reactivity towards glucose oxidase is dictated only by redox potential, PQQ analogues with unchanged o-quinone group are expected to react with glucose oxidase in a similar way as native PQQ. Several PQQ derivatives with alkyl residues at N_1 and respectivily C_3 position of the pyrrole ring were investigated with respect to their reactivity to function as electron acceptor for glucose oxidase (Table 2). It was shown that the N_1-alkyl derivatives of PQQ react slowly with glucose oxidase. The second order rate constant for the oxidation of reduced glucose oxidase by N1-methyl-PQQ and N_1-ethyl-PQQ at pH 3.0 is approximately 102 fold smaller than that for PQQ. Under the same conditions 3-methyl-PQQ oxidizes reduced glucose oxidase with a rate comparable with that for PQQ. These results indicate that structural modification of the PQQ molecule may occur during its reaction with reduced glucose oxidase and is responsible for PQQ reactivity.

To obtain a better insight in the mechanism of the reaction of PQQ with reduced glucose oxidase, the reaction of the enzyme with other organic compounds with a quinoid structure such as 2,6-dichloroindophenol (Cl_2Ind) and p-benzoquinone was examined (Fig. 3). p-Benzoquinone reacts with high rates with glucose oxidase over the entire pH range 2.4—8.0, suggesting that p-benzoquinone is able to react

in protonated and unprotonated form with the reduced FAD within glucose oxidase. Cl_2Ind shows a similar pH dependence behaviour as PQQ. This points to a similar reaction mechanism for the reaction of PQQ and Cl_2Ind with reduced glucose oxidase in spite of an apparent structural difference, Cl_2Ind being a para-quinonimine and PQQ an ortho-quinone.

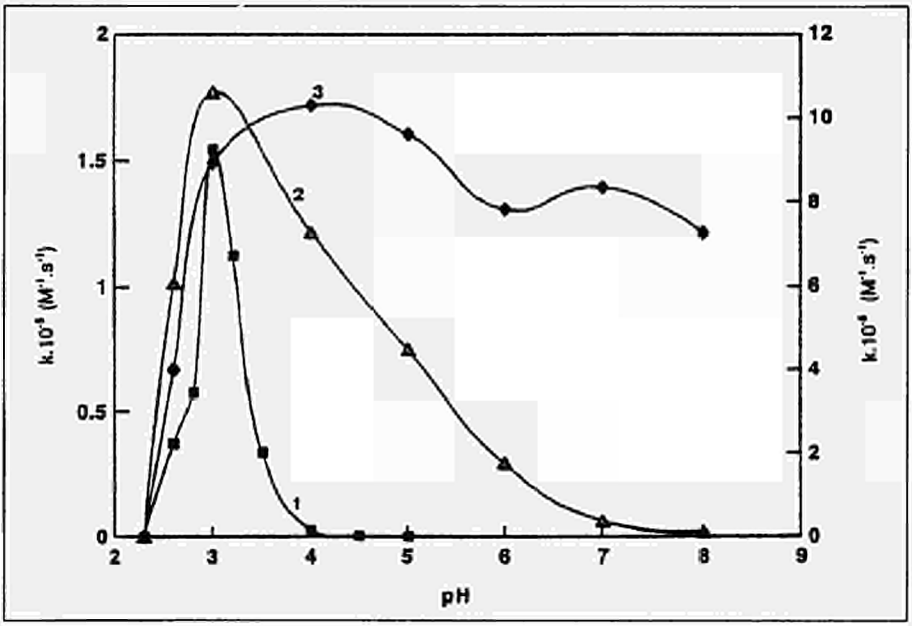


Figure 3. pH-dependence of the second-order rate constant for: (1) pyrrolo-quinoline quinone; (2) 2,6-dichloroindophenol; (3) p-benzoquinone (on right axis)

The explanation for the similarity in reactivity may be that the reactive structure of PQQ is the para-quinoid form caused by base-catalysed tautomerization. However, Cl_2Ind still reacts with glucose oxidase above pH 5.0. Beside the structure of the electron acceptor molecule its charge affects the interaction with the protein environment and influences the pH-dependence of the reaction. For PQQ the carboxylic groups with pK_a values ranging from 5.0 to 2.0 affect the onset and rate of the reaction. Obviously, the carboxylic groups with a pK_a of around 5.0 must be protonated before, presumably, a charge transfer-complex formation and subsequent catalysis can take place.

More detailed information on the interaction of PQQ and glucose oxidase have been obtained from separate spectral studies between PQQ and FMN in comparison to PQQ and glucose oxidase as holoenzyme.

Information about the interaction between PQQ and FMN were obtained from absorption, fluorescence and ^{1}H-NMR spectroscopy. The spectral data suggest the formation of a charge-transfer complex between the two molecules due to interactions between their aromatic ring systems. Also hydrophobic associations between these aromatic ring systems may contribute to the complex formation. The complexation between PQQ and FMN is characterized by a dissociation constant K_d = 1.92 μM, determined by difference spectral titration. The observation that PQQ

is a strong complexer with respect to FMN points to importance of this association for its reaction with FAD containing glucose oxidase.

Using absorption and fluorescence spectroscopy to study the interaction between PQQ and glucose oxidase, a strong association between the two molecules in the acidic pH range was observed. Spectral data suggest that there are at least two possible ways of interaction of PQQ with glucose oxidase, one through hydrogen bonding at the outer layer of the enzyme molecule, and another through hydrophobic binding and complexation with the flavin part of FAD and several amino acid side chains in the particular environment of the FAD binding site.

By quantifying the interaction between PQQ and glucose oxidase the following results were obtained: at pH 3.0 one dimer molecule of glucose oxidase binds one molecule of PQQ with a dissociation constant K_d = 0.82 μM and 2 PQQ molecules with K_d = 15.2 μM; at pH 4.0 one molecule of glucose oxidase binds one molecule PQQ with a dissociation constant K_d = 30.6 μM.

Based on these observations it can be concluded that the reaction of PQQ with reduced glucose oxidase is determined by the specific properties of the three molecules involved, i.e. the reduced flavin part of FAD, glucose oxidase apoprotein and PQQ. As a result the reaction between PQQ as electron acceptor and reduced glucose oxidase has three distinct steps:

a)Diffusion of the PQQ molecule within the protein (controlled by electrostatic interaction).
b)Tautomerization of the PQQ molecule to achieve the para-quinonimine active form.
c)Oxidation of the reduced FAD within the enzyme.

b) Poly-viologens as electron acceptors.

Another class of artificial electron acceptors for glucose oxidase are polymers based on 4,4'-bipyridyl (polyviologens). Polyviologens were synthesized by a reaction between equimolecular amounts of 4,4'-bipyridyl and α,α'-dibromo-o- and p-xylene. The polymeric products were purified and characterized by IR, UV and 1H-NMR spectroscopy. The polymers were investigated spectrophotometrically as electron acceptors for glucose oxidase by monitoring the increase of the absorption at 632 nm, characteristic for semiquinone formation. The reaction takes place with a maximum rate at pH 7.5. The turnover number k_{cat} and K_m values for the polyviologens were determined. It was observed that, in time, the semiquinone participates in a side reaction that results in a dimer formation due to the folded structure of the polymer, characterized by an absorption maximum at 535 nm. As a result of this dimerization the polymer does not mediate the electron transfer between the enzyme and the electrode in an electrochemical cell when the electrolyte solution contains the natural enzyme, solubilized mediator and glucose.

To eliminate this side reaction N-aminopropyl-poly-o-xylylviologen was synthesized, a polymer with a terminal amino group that might be incorporated within glucose oxidase.

The polymer was attached to glucose oxidase by covalent binding to the protein part of the enzyme by amide links between the polymer amino group and enzyme carboxylic groups by the carbodiimid method. Around 6 molecules of polymer per molecule glucose oxidase were bound. This modified glucose oxidase directly transfers electrons to a glassy carbon electrode. As a result a glucose concentration

dependent current flows in an electrochemical cell when the electrolytic solution contains the polymer modified glucose oxidase.

2. *NADH oxidase from* Thermus thermophilus

NADH oxidase from *Thermus thermophilus* is a flavoprotein that catalyses the oxidation of NADH by molecular oxygen with the formation of NAD^+ and H_2O_2 that can use both FMN and FAD as coenzyme, that are both unusually loosely bound to the enzyme. NADH oxidase catalyses the regeneration of NAD^+. The enzyme is of potential interest for de development of biosensors, since NAD^+ dependent analytically important reactions catalysed by dehydrogenases can be both detected by amperometric signal and amplified by NAD^+ regeneration.
It could be demonstrated that the enzyme can catalyse NADH oxidation with redox compounds like PQQ and 2,6-dichloroindophenol, even in the absence of a flavocoenzyme.

For PQQ and Cl_2Ind under anaerobic conditions in the presence of 210 µM NADH the following kinetic parameters were determined: k_{cat} 17 s-1 and K_m 22.3 µM for PQQ and k_{cat} 31 s^{-1} and K_m 4.2 µM for Cl_2Ind.

FAD reacts also as electron acceptor for NADH oxidase under anaerobic conditions: k_{cat} 4.8 s^{-1} and K_m 53.1 µM. Surprisingly, FAD is a less favourable electron acceptor for this NADH oxidase than PQQ and Cl_2Ind.

3. *NADH oxidase from* Thermus aquaticus

NADH oxidase from *Thermus aquaticus* is a strictly FAD dependent flavoprotein catalysing the same reaction as mentioned for NADH oxidase from *Thermus thermophilus*. The FAD is also very loosely bound for this NADH oxidase.

Investigations on the development of a NADH oxidase based amperometric system for the detection of NADH and for the amplification of the signal for NAD dependent by dehydrogenases catalysed reactions of analytical interest were continued.

It is a general experience that the direct oxidation of NADH causes poisening of the electrode surface. This therefore needs to be kept to a minimum for the biosensor to function effectivily. Therefore the operating potential of the electrode has to be below the oxidation potential of NADH, for example, at a bare glassy carbon electrode which is +560 mV.

The electrochemistry of three quasi-reversible water-soluble ruthenium complexes exhibiting low oxidation potentials was examined. These were:

Ruthenium hexamine trichloride: $(Ru(NH_3)_6)Cl_3$

Ruthenium red (RR): $((NH_3)_5Ru\text{-}O\text{-}Ru(NH_3)_4\text{-}O\text{-}Ru(NH_3)_5)Cl_6$

Ammonium hexachlororuthenate: $(NH_4)_2(RuCl_6)$

Ruthenium hexamine trichloride and ammonium hexachlororuthenate exhibit a single set of redox peaks corresponding to the Ru(II/III) and Ru(III/IV) couples respectivily (for Ru(II/III):
$E_{1/2}$ = -175 mV, ΔEp = 80 mV at 50 mV/s; for Ru(III/IV): $E_{1/2}$ = 80 mV, ΔE_p = 155 mV at 50 mV/s).

The $RuCl_6^{3-}$ once formed undergoes hydration over a period of time with the rate of substitution for the first Cl^- of the order of seconds. The structure of the species undergoing mediation is therefore probably $(RuCl_5H_2O)^{2-}$.

Ruthenium red is a trinuclear amine complex of the form (Ru^{III}-Ru^{IV}-Ru^{III}). It has been shown to undergo a series of one-electron-transfer reactions leading eventually to the formation of (Ru^{V}-Ru^{V}-Ru^{V}). The first redox couple of ruthenium red was examined ($E_{1/2}$ = -190 mV, ΔE_p = 60 mV at 50 mV/s). This corresponds to the formation of ruthenium brown (Ru^{IV}-Ru^{III}-Ru^{IV}).

The first experiments concerned the possible catalytic activity to NADH by each mediator and by unbound FAD. Amperometric investigations were carried out at the foreseen working potential of the biosensor. Therefore FAD, ruthenium hexamine trichloride and ruthenium red were used at 0 mV while ammonium hexachlororuthenate was used at + 150 mV. Mediators were present at concentrations of 0.61 mM while the FAD concentration was 0.41 mM. The response to a NADH concentration of 0.82 mM was measured in duplicate. None of the mediators when used alone gave an increase in the rate of NADH oxidation. However, in the presence of FAD all the three mediators showed some degree of electrocatalysis:

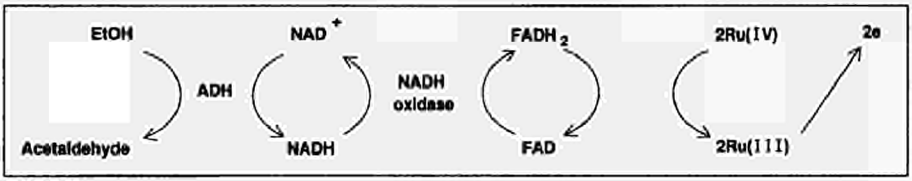
Potential (mV)	*Oxidant*	*Current (nA)*
0	None	18
	FAD	–
	RR	–
	$Ru(NH_3)_6{}^{3+}$	–
	RR/FAD	40
	$Ru(NH_3)_6{}^{3+}$/FAD	220
150	None	79
	FAD	–
	$RuCl_6{}^{2-}$	–
	$RuCl_6{}^{2-}$/FAD	108

The electrocatalytic current produced by $Ru(NH_3)_6{}^{3+}$/FAD is considerably greater than the uncatalysed background value for the direct oxidation of NADH. Ruthenium hexamine trichloride was therefore considered unsuitable for use with NADH oxidase.

Ruthenium red and ammonium hexachlororuthenate were therefore examined for coupling with NADH oxidase.
The coupling with the enzyme was performed in the presence of oxygen and the electrode was poised at the oxidation potential of each mediator. NADH was initially added to the cell containing mediator and FAD. After the electrocatalytic oxidation had reached a steady state, NADH oxidase at a concentration of 4 µg/ml was added. Both mediators produced a significant current increase:

Potential (mV)	*Oxidant*	*Current (nA)*
0	RR/FAD/Enzyme	432
150	$RuCl_6{}^{2-}$/FAD/Enzyme	437

The same concentration of mediator and FAD gave no response to approximately 770 mM H_2O_2. This indicated that the observed increase in current in the presence of both mediators is not due to electrocatalysis of oxidation but instead to the ruthenium complexes replacing oxygen as the oxidant to the reduced form of NADH oxidase. The removal of oxygen from the electrochemical cell was found to cause an increase in the response suggesting that full substitution of the oxygen by the mediators had not taken place. However, since the response in the presence of mediators was considered to be sufficiently high, the construction of an NADH oxidase based amplification system operating at a lower potential by mediating electron transfer through the $RuCl_6^{2-}/RuCl_6^{3-}$ the redox couple was investigated. The viability of the following amplification scheme was initially tested using NADH as the trigger for the reaction.



Scheme of end of part NADH oxidase from Thermus aquaticus

This was carried out in order to demonstrate the scale of any possible electrocatalytic effects. At relatively low NADH concentrations (less than 12 μM) used no current was observed in the absence of NADH oxidase. However, when NADH oxidase was introduced in the system a value of 358 nA/μM was observed. The amplification system was then extended for the determination of ethanol as illustrated in the scheme above. The NADH oxidase was immobilized on the electrode using both carbodiimide activation of glassy carbon and physical adsorption following cross-linking with glutaraldehyde. Lineweaver-Burk plots were obtained for each electrode treatment using $RuCl_6^{-2}$ in bulk solution to provide the analytical signal. The values obtained for the apparent K_m and V_{max} were 172 μM and 0.77 μA, and 66 μM and 0.22 μA for glutaraldehyde and carbodiimide immobilized NADH oxidase respectively. These K_m values for the electrode immobilized NADH oxidase are considerably higher than the solution value. This is probably due to the unstirred Nernst diffusion layer at the electrode, the possible diffusion barrier presented by any immobilized inactive enzyme and the lower rate constant for enzyme oxidation by the mediator. Comparison between the two immobilization procedures suggests that the glutaraldehyde method produces a greater mass of enzyme present on the electrode. Crosslinking was therefore chosen as the preferred form of immobilization. The response of the electrode to ethanol in the presence of cross-linked NADH oxidase was carried out after alcohol dehydrogenase deposited and crossîlinked. The electrode was found to respond linearly to an ethanol concentration of 20—200 μM.

In parallel with these experiments, it was discovered that NADH measurement could be performed in the absence of NADH oxidase and a water-soluble mediator. The conducting polymer poly(indole-5-carboxylic acid), in film-form on a glassy carbon disc electrode, was observed to be inherently electrocatalytic towards NADH. A similar phenomenon was observed towards ascorbate. The electrocatalytic effect of poly(indole-carboxylic acid) deposited on the electrode with respect to NADH and ascorbate is illustrated by the cyclic voltammograms of both analytes at bare and modified electrodes (Fig. 4 and Fig. 5).

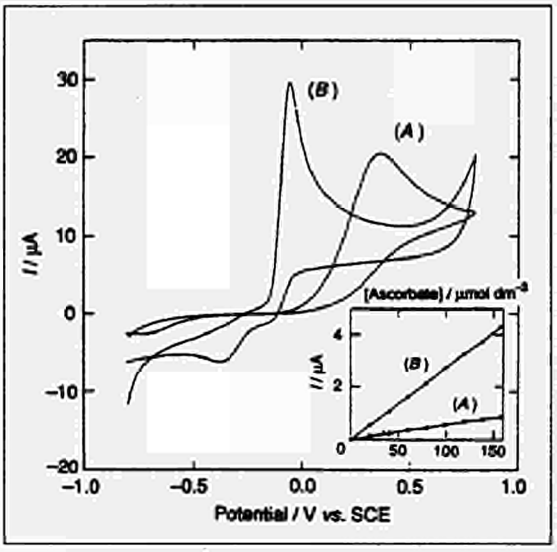


Fig. 4. Cyclic voltammogram of 2 mmol dm^{-3} ascorbate in Tris-HCl, pH 8.8 at bare (A) and polymer-modified electrode (B). Scan rate: 20 mV s^{-1}. Inset: calibration of ascorbate at 0 mV at bare (A) and modified (B) electrode.

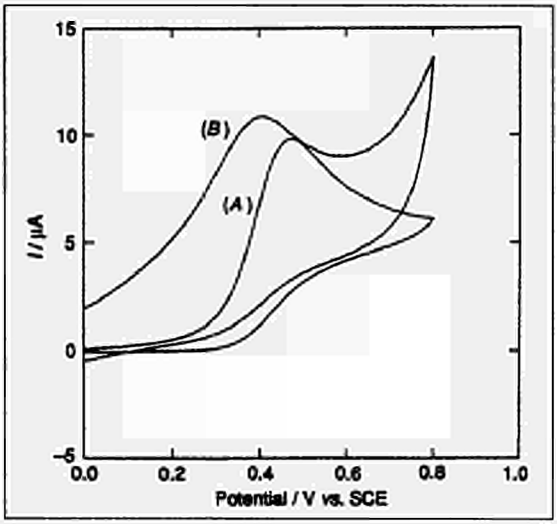


Fig. 5. Cyclic voltammogram of 2 mmol dm^{-3} NADH in Tris-HCl, pH 8.8 at bare (A) and polymer-modified electrode (B). Scan rate: 20 mV s^{-1}

In both cases polymer deposition also caused an increase in the effective surface area, as is evident in the voltammograms by increased non-faradaic currents. The shift in oxidation potential is greater for ascorbate than for NADH. At this stage the reason for this is unclear and is perhaps related to the particular mechanism for the NADH oxidation. It would appear that the polymer is apparently electrocatalytic without acting as an electron-transfer mediator. The reasons for suggesting this are that in buffer solution the polymer film shows only an irreversible reduction peak (-400 mV vs standard calomel electrode, SCE), whereas to mediate

oxidation of the analyte, the polymer would need to be reoxidized at the electrode. Hence, this catalytic effect can perhaps be viewed as analogous to electrode modifications such as platinization, in which the oxidation potential of an analyte is lowered without the action of a redox mediating species. Ascorbate and NADH were determined by amperometry at 0 and + 450 mV vs SCE respectively. In each case calibrations were performed before and after electrode modification. As illustrated in Fig. 6 and the inset to Fig. 4, the increase in sensitivity to both analytes is similar. In addition to this, polymer modification prevented electrode poisoning during NADH oxidation, producing steady state rather than peaked responses and increasing the linear range, as can be seen by comparison with the inset to Fig. 6 . A potential application of NADH determination by this principle may be NAD dependent ethanol analysis in blood or serum using alcohol dehydrogenase immobilized by crossîlinking with glutaraldehyde onto such a polymer modified electrode. Precautions with respect to diminution of signals of interfering compounds, for example, of ascorbate should be considered.

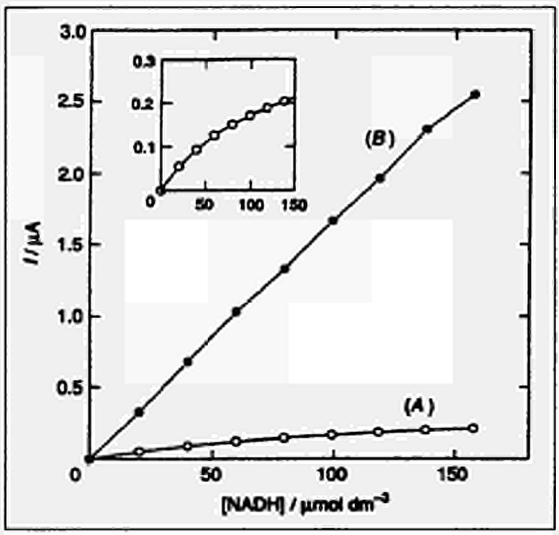


Fig. 6. Calibration of NADH at 450 mV vs SCE at bare (A) and polymer-modified (B) electrode. Bare electrode calibration shown with expanded y axis.

D. Enzymological studies

1. The flavin domain of cytochrom b2 (L-lactate dehydrogenase) from Saccharomyces cerevisiae

A clone of the gene encoding the flavin domain of cytochrome b_2 (NAD independent L-lactate dehydrogenase) from *Saccharomyces cerevisiae* was obtained as an insert in a plasmid in *Escherichia coli* (supplied by Drs S.K. Chapman and C. Brunt, University of Edinburgh).

In the last report a procedure was described for the purification of this flavin domain. Now a reliable procedure has been developed for the reversible removal of FMN from the recombinant flavin domain of cytochrome b_2 (flavocytochrome b_2). When treated on a column of Sephadex G-25 at moderate low pH, e.g. 3.75, 97% of the FMN was removed from the enzyme. Addition of FMN at pH 7 to the apo-enzyme restored 53% of the activity. The apo-enzyme was stable for at least 16 h, making feasible reconstitution experiments with FMN analogues.

2. *NADH oxidase from* Thermus aquaticus

NADH oxidase was purified from *Thermus aquaticus* cell paste. Cell-free extracts were subjected to chromatography on DEAE-cellulose, phenyl-Sepharose and Sephacryl S200 according to published procedures, withn a final fractionation by preparative PAGE. The specific activity of the enzyme in a spectrophotometric assay for NADH oxidation was 13 and the enzyme was judged to be pure by PAGE and SDS-PAGE (M = 54,000). The absorption spectrum of the preparation had maxima at 274 nm, 384 nm and 450 nm with an A274/A450 ratio of 8.1, similar to that of other purified flavoenzymes. In agreement with published work, however, flavin and protein analyses suggest that the preparation contains about 0.3 FAD/subunit. The enzyme catalyses only a slow rate of NADH oxidation with oxygen as electron acceptor unless FAD is added to the reaction mixture. However, the enzyme is highly active with the electron acceptors 2,6-dichlorophenol-indophenol and ferricyanide ion even in the absence of exogenous flavin (FAD, FMN, or riboflavin), suggesting that the rate-limiting step in the assay with oxygen is transfer of electrons from the enzyme bound flavin hydroquinone. It is likely that exogenous flavin mediates electron transfer from the hydroquinone of the enzymeîbound flavin to oxygen. Also a marked inhibition by NADH was observed. The inhibition is observed with even the best quality NADH available commercially, but the possibility that it is due to contaminants or breakdown products in the NADH, as observed with certain other NADH dependent dehydrogenases, can not be excluded. Structural elucidation is planned. This is justified, because the NADH oxidases from *T. thermophilus* and *T. aquaticus* are probably quite different enzymes given their distinct flavocoenzyme preference. The fact that also a pure homogeneous preparation of this NADH oxidase could be obtained has allowed the initiation of molecular biological investigations on the enzyme. The aim of this investigations is to clone and express this enzyme in *E. coli.*

The N-terminal amino acid sequence has been determined:

Ala-Asp-Ser-Thr-Gly-Glu-His-Phe

Further investigations have been plannned:

a) Generation of a list of probes from the amino acid sequence
b) Isolation of pure DNA from *T. aquaticus*
c) Digestion of the DNA with various restriction enzymes
d) Separation of fragments on agarose gels and transfer to hybridization membranes
e) Hybridization with selected probes and determination of the hybridization conditions
f) Selection of hybridizing fragments and determination of their size

3. N^6-*(2-aminoethyl)-FAD as coenzyme for apo-NADH oxidase from* Thermus thermophilus *and* Thermus aquaticus

The coenzyme properties of N^6-(2-aminoethyl)-FAD in comparison to FAD at 30°C and 50°C have been studied with respect to the apo-NADH oxidase from *Thermus thermophilus* (FMN and FAD dependent) and from *Thermus aquaticus* (strictly FAD dependent) to both of which the flavin coenzymes bind very loosely. The data have been summarized in Table 3.

Both FAD and N^6-(2-aminoethyl)-FAD are accepted as comparable coenzymes for the apo-form of both NADH oxidases. With respect to structural interaction, the increased K_m values for N^6-(2-aminoethyl)-FAD reflect a sterically somewhat hindered alignment of the N^6-2-aminoethyl)-adenine fragment in the adenine part of the FAD binding site for both apo-enzymes. Preliminary kinetic results with apo-NADH oxidase from *Thermus thermophilus* point to a more favourable acceptance of N(1)-(2-aminoethyl)-FAD, similar to that of FAD. This is in agreement with the observation from the structure of this NADH oxidase that the N(1)-position of the adenine of FAD is sterically better accessible from the outside than the N^6-position. Raising the temperature from 30°C to 50°C may cause a slight structural change in the apo-NADH oxidase from *Thermus aquaticus* leading to a somewhat stronger interaction of N^6-(2-aminoethyl)-FAD.

Since the V_{max} values remain constant at both temperatures, it apparently also leads to a hindered electron transport from NADH to the isoalloxazine part of this FAD analogue.

These kinetic results consolidate positive perspectives for the preparation of enzymatically active covalent NADH-oxidase-FAD-analogue adducts to overcome the disadvantage of a very weak bonding of FAD to both native NADH oxidases.

Table 3. Comparison of K_m and Vmax for N^6(-aminoethyl)-FAD and FAD with respect to apo-NADH oxidase from *Thermus thermophilus* and *Thermus aquaticus* at 30°C and 50°C

	K_m (μM)		V_{max} (U/mg)	
T thermophilus	*30°C*	*50°C*	*30°C*	*50°C*
FAD	0.054	0.055	44	55
N^6-(2AE)-FAD	0.255	0.264	37	83
Conditions: 42 μM potassium phosphate, pH 7.2, 0.165 μM NADH.				
T. aquaticus				
FAD	0.043	0.043	5.2	9.3
N^6-(2AE)-FAD	0.220	0.133	7.2	7.5

Conditions: 20μM Tris-HCl 7.4, 0.176 μM NADH; U: International Unit (Apo-NADH oxidases were purified to homogeneity).

E. Screening of thermophile or thermotolerant bacteria exhibiting thermostable oxidase activity of interest for the development of biosensor systems

Although many types of thermophilic or thermotolerant aerobic microorganisms exist, the search for the presence of thermostable oxidases of potential for the development of biosensor systems is a neglected field in applied microbiology.

Efforts have been undertaken to enrich and purify thermophilic bacteria from natural environment (soil, sediments, combust, bark and straw) and industrial plants (hot production waters of a local dairy and sugar factory), but also from the German Collection of Microorganisms (DSM), to screen for oxidase activity during

growth on different carbon sources. With respect to a potential application in biosensors, a screening was continued with emphasis on glucose-, lactate- and glycerol- and L-amino acid oxidase. A reliable qualitative test system has been developed for agar plates to detect H_2O_2 formed by the oxidase screened for. The test is based on the by *in vitro* added peroxidase catalysed conversion of the chromogen ABTS reacting with H_2O_2 (green colour development) in the presence of 3-amino-1,2,4-Triazol, inhibitor for *in vivo* interfering H_2O_2 splitting catalase. From enrichment cultures inoculated agar media, supplemented with the appropriate substrates and the inhibitor, were first incubated at 60°C till colonies appeared. Then an agar overlay with peroxidase and ABTS was deposited and the incubation at 60°C continued for one day. After one day at this temperature the thermolable peroxidase was still satisfactory active to develop green colouring where H_2O_2 was present on the plate.

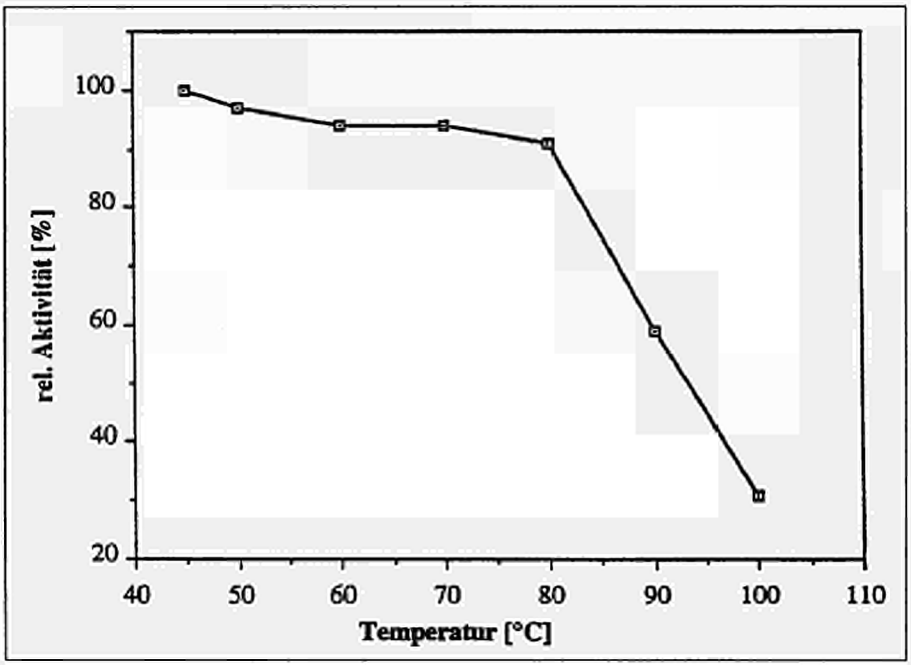


Figure 7. Thermostability of L-histidine oxidase from thermophile Bacillus *species. Incubation conditions: 0.05 ml partly purified and concentrated enzyme (0.9 mg/ml) in 0.025 mM potassium phosphate, pH 6.5, in the temperature range 45 — 100°C with incubation time 10 min. In a standard assay the rest activities at the different temperatures were percentually related to the rest activity at 45°C*

An alternative to this test, trapping by oxidase activity generated H_2O_2 with cerium chloride ($CeCl_3$) and subsequent conversion of cerium perhydroxide to a brown-black precipitate by reaction with 3,3'-diaminobenzidine, will be investigated. If in agar medium H_2O_2 might be satisfactory detected with such test, the insertion of a peroxidase step can be omitted. Around 5000 colonies have been screened for oxidase activity with the ABTS method. After enrichment, 15 glucose-, 25 lactate-, 9 glycerol- and 3 L-amino acid oxidase positive colonies were isolated. Only 2 ther-

mophile bacteria were obtained, that could be further cultivated in liquid medium: Strain RS-1 from combust, pleomorphic and presumably a *Corynebacterium* with lactate oxidase activity, and strain NO-2 from a sediment, endospore forming rod and presumably a *Bacillus species* with L-amino acid oxidase activity with, surprisingly, substrate specificity for L-histidine and L-phenylalanine having both an aromatic ring in the side chain. Both oxidases were formed under cultivation conditions with slight oxygen limitaton after the transition from the logarithmic to the stationary growth phase. The oxidase from *Bacillus species* is an intracellular enzyme and, presumably, associated with endospore formation. The L-amino acid oxidase from the *Bacillus species* was partly purified by anion exchange chromatography and preliminary characterized. After native PAGE and activity staining with L-histidine as substrate it was demonstrated that the enzyme has a molecular weight of around 52 kDa. The enzyme shows preference for L-histidine compared to L-phenylalanine in a standard assay (activity 30% of that for L-histidine) and should be called L-histidine oxidase. This seems to be a new oxidase since L-phenylalanine oxidase from *Pseudomonas species* P-501 shows a very minor substrate specificity for L-histidine (H.Koyama, *J.Biochem.* **92**, 1235-1240 (1982)). For the new L-histidine oxidase the pH and the temperature optima were 6.0-6.5 and 45-50°C respectivily. This oxidase is thermostable as is illustrated in Fig. 6. If by pH shifting the substrate preference might be changed to L-phenylalanine, will be investigated. Using L-phenylalanine oxidase from Pseudomonas species for determining L-phenylalanine by oxidative deamination producing H_2O_2 is severely complicated by the simultaneously catalysed oxidative decarboxylation producing CO_2. Now it should be investigated, if the new L-histidine oxidase catalyses only the for amperometric biosensors important oxidative deamination. Efforts should be undertaken to clone L-histidine oxidase with the aim to produce relatively large quantities for further characterization. For application, the high thermostability and the satisfactory activity at moderate temperatures (30-45°C) are favourable properties for application purposes. Unfortunately, conditions for the cultivation of strain RS-1 at such a scale that enough lactate oxidase could be partially purified for a similar preliminary characterization have not been found yet. In the last part of the project a screening program will be pursued with emphasis on thermotolerant microorganisms including fungi and yeasts.

MAJOR SCIENTIFIC BREAKTHROUGHS

1. Development of a synthesis strategy to prepare N^6-(carboxyalkyl)-FAD by diphosphate coupling of FMN and the heptylester of N^6-(carboxyalkyl)-AMP and subsequent enzymatic deprotection of the carboxyl group by lipase as exemplified for N^6-(6-carboxyhexyl)-FAD.

2. Preparation of a covalent adduct between N^6-(6-carboxyalkyl)-FAD and apo-D-amino acid oxidase from pig kidney containing 1 covalently bound FAD analogue molecule/subunit (determined by electro-spray MS) with an activity comparable to that of the native holo-enzyme which points to a similar structural position as native FAD (The adduct could be satisfactory crystallized).

3. Similar acceptance of FAD and N^6-(2-aminoethyl)-FAD as coenzyme for the apoform of NADH oxidase from *Thermus thermophilus* and *Thermus aquaticus*.

4. Application of ammonium hexachlororuthenate and N-aminopropyl-poly-o-xylylviologen as new redox mediators in enzymatic biosensor systems for the quantitative determination of respectivily ethanol and glucose.

5. Discovery of a specific mediatorless electrocatalysis at a conducting polymer electrode with respect to NADH and ascorbate

6. Discovery of a new oxidase in a thermophile *Bacillus species*, L-histidine oxidase with subtrate specificity for L-histidine and L-phenylalanine.

The results obtained sofar may lead to a better understanding of the structural aspects of electron transfer to electrodes from FMN and FAD dependent oxidases of interest for the development of amperometric biosensors and how to design such oxidases, chemically by covalent attachment of redox mediators and/or by site-directed mutagenesis, to improve their properties as element of biosensor devices.

MAJOR COOPERATIVE LINKS

Third meeting for participants of the project at the Agrotechnological Research Institute, Wageningen (NL), February 7-8, 1994.

Exchange of both types of NADH oxidases between the Biotechnology Centre of the Cranfield Institute of Technology and the Gesellschaft für Biotechnologische Forschung with intention to prepare a common publication in the last year of the project.

Supply of cell material of *Thermus aquaticus* by the Biotechnology Centre of the Cranfield Institute of Technology to the University College Dublin.

Organization of the intermediate sectorial meeting on Biotransformations in the BRIDGE programme at the Gesellschaft für Biotechnologische Forschung, Braunschweig (D), May 4-6, 1993.

PUBLICATIONS

Joint publications

A.F. Bückmann, H. Erdmann, M. Pietzsch, J.M. Hall and J.V. Bannister, Synthesis and coenzyme activity with respect to apo-NADH oxidase from *Thermus thermophilus* and *Thermus aquaticus*, Proceedings 11th International Symposium on Flavins and Flavoproteins 1993, Nagoya, Japan (in press).

Individual publications

A.F. Bückmann, V. Wray and H.C. van der Plas, Simultaneous conversion of N(1)-(2-aminoethyl)-adenosine to N^6-(2-aminoethyl)-adenosine and tricyclic 1.N6-ethanoadenosine under mild aqueous conditions, Heterocycles (submitted for publication)

C. Boeriu and C. van Dijk, The reaction of glucose oxidase with one- and two-electron acceptors, Proceedings of the National Conference on Sensor Technology 1994, Enschede, Netherlands, P,V, Lambeck, Ed. p. 247 — 251

M. Somasundrum and J.V. Bannister, Mediatorless electrocatalysis at a conducting polymer electrode: Application to ascorbate and NADH measurement, J. Chem. Soc., Chem. Commun. **21**, 1629 — 1631 (1993)

M. Somasundrum, J.M. Hall and J.V. Bannister, Amperometric NADH determination via both direct and mediated electron transfer by NADH oxidase from *Thermus aquaticus*, J. Chem. Soc., Chem. Commun. (submitted for publication)

New types of redox enzymes for the production of chiral synthons: basic research, functionalisation and application (BIOT CT-900157)

COORDINATOR:
H. GÜNTHER / H. SIMON, Org. Chemie & Biochemie, TU München, Garching, D
PARTICIPANTS:
W. SOMERS, Netherlands Organiz. for Appl. Sci. Res. (TNO), Zeist, NL
J.A. DUINE, Delft University of Technol., Fac. Chem. Technol. Delft, NL
I. MOURA, Univers. Nova de Lisboa Fac. Ciências e Technol. Lisboa, P
ASSOCIATED PARTICIPANT (to TNO):
C. van DIJK, ATO-DLO Wageningen, NL

BACKGROUND INFORMATION

The use of the 850 known oxidoreductases, 350 of them pyridine nucleotide (PN) dependent, for the production of chiral synthons will be economic only if a simple and inexpensive source of reduction equivalents is available.

OBJECTIVES AND PRIMARY APPROACHES

For the production of chiral synthons the expensive PNs have to be regenerated or omitted. In order to avoid PNs, the natural redox mediators, cheap artificial redox mediators have to be developed and enzymes which accept them have to be detected and characterised.

RESULTS AND DISCUSSION

The group in Munich:
a) Detected a very effective regeneration system for all 4 pyridine nucleotides by artificial mediator accepting pyridine nucleotide oxidoreductases (AMAPORs) using various artificial mediators.
b) Partially characterised the AMAPOR activities of *Clostridium thermoaceticum*.
c) Isolated and partially characterised three 3-oxo ester oxidoreductases and a 3-oxo acid reductase.
d) Applied the AMAPOR system for the preparation of different chiral products with high added value.
e) Characterised in cooperation with the groups in Delft and Lisboa two different, reversible aldehyde oxidoreductases from *Clostridium formicoaceticum*, one containing tungsten and one molybdenum.
f) Worked on the use of new artificial mediators.

We detected high and rather stable AMAPOR activities in crude extracts of *C. thermoaceticum* for NAD(P)H and $NAD(P)^+$ regeneration using artificial redox mediators. Crude extracts showed various proteins with NADP-dependent AMAPOR activity but only one NAD-dependent activity. AMAPOR containing protein fractions had spec. act. of up to 1500 U/mg protein for NADPH regeneration and molecular masses of 300 and 600 kD with 42 and 56 kD subunits and 200 and 400 kD with 42 and 76 kD subunits, respectively. A decrease of activity by ion exchange chromatography could be reversed by adding FAD (3 µM saturating), FMN had no effect. We also purified and characterised 1 enzyme from *Clostridium*

kluyveri and 2 enzymes from *Clostridium tyrobutyricum* reducing 3-oxo esters to [S]-3-hydroxy esters. An enzyme from *C. tyrobutyricum*, which reduces 3-oxo acids seems to be the first such enzyme from procaryotes. All reduce 3-oxo acid moieties to [S]-3-hydroxy acid moieties with ee-values >98%. The two enzymes from *C. tyrobutyricum* reducing the esters are NAD- the other enzymes NADP-dependent. The newly detected reversible 3-oxo acid reductase has a spec. act. of 1400 U/mg protein and a molecular mass of 40-43 kD and does not contain FAD or FMN. The four enzymes were partially sequenced starting at the amino end.

With these enzymes 3-oxo esters and 3-oxo acids were reduced using various regenerating systems for the pyridine nucleotides. From racemic 3-hydroxybutyrate the [S]-enantiomer can be selectively dehydrogenated by the 3-oxo acid reductase leaving the R-enantiomer unreacted. The enzyme has been applied also in the form of rethawed whole cells.

The preparative regeneration of NAD(P)H and $NADP^+$ was performed in electrochemical cells using crude extracts of *C. thermoaceticum* and an artificial mediator of suitable redox potential. Since anaerobic extracts of *C. thermoaceticum* also possess an NADP dependent formate dehydrogenase the preparative NADPH regeneration can be achieved with significantly lower apparative expenses using formate as electron donor instead of the cathode of an electrochemical cell. With formate as electron donor [S]- 3-hydroxybutyrate, [S]-3-hydroxyvalerate and [S]-3-hydroxycaproate and [S]- 3-hydroxyisocaproate were prepared with ee >98% and 95%-100% yields, respectively. The corresponding esters have been prepared.

Using enzyme relay systems with AMAPOR of crude extracts or cell suspensions of *C. thermoaceticum* and pyridine nucleotide dependent oxidoreductases in electrochemical cells we synthesised on a preparative scale:

(i) [2R,3S]-isocitrate by reductive carboxylation of 2-oxoglutarate (value added factor 100).
(ii) The opposite enantiomer [2S,3R]- isocitrate with productivity numbers of up to $1.4*10^4$ from racemic isocitrate by oxidative decarboxylation of the [2R,3S]-enantiomer (i + ii catalysed by isocitrate dehydrogenase).
(iii) 6-phosphogluconate from glucose 6-phosphate by glucose 6-phosphate dehydrogenase.
(iv) ribulose 5-phosphate by the decarboxylating dehydrogenation of 6-phosphogluconate catalysed by 6-phosphogluconate dehydrogenase.

Three commercially not available viologens were synthesised:

(i) 1,1'-2-(1,3-dioxolane-2"-yl)ethyl-4,4'-bipyridinium salt embodies easily available aldehyde groups in the dioxolane moiety, possibly useful in immobilisation experiments,
(ii) 1,1'-(p-phenylsulphonato)methyl- 4,4'-bipyridinium salt, neutral at physiological pH in oxidised and negatively charged in reduced state.
(iii) 1-CH_3-1'-[$HOOCCH_2$-(O-CH_2-CH_2)$_n$-O -CH_2CO-NH-(CH_2)$_3$]-bipyridinium salt, with a functionalised polyethylene glycol spacer.

The Delft group investigated the properties of different classes of dye-linked alcohol and aldehyde dehydrogenases of interest for application in amperometric biosensors, test strips, electrochemical production cells and kinetic resolution processes for biotechnological production of homochiral fine chemicals.

Quinohaemoprotein alcohol dehydrogenase, QH-EDH, from *Comamonas testosteroni* was found to have excellent enantioselective properties for the oxidation of chiral solketal, 2,2-dimethyl-4-hydroxymethyl-1,3-dioxolane. Since both

homochiral S- and R-solketal are highly attractive starting compounds for the production of chiral pharmaceuticals, including the β-blockers propranolol and atenolol, the merits of this enzyme for biotechnological applications have been explored. QH-EDH is isolated as the inactive haem c-containing PQQ free apo enzyme from ethanol grown *C. testosteroni*. Upon addition of PQQ, 2,7,9-tricarboxy -1H-pyrrolo[2,3-F]quinoline-4,5-dione, an active alcohol dehydrogenase is obtained with spec. act. 28 U/mg on n-butanol using ferricyanide as an electron acceptor for the 72 kDa monomeric protein. Irregular yields and activities could be contributed to the presence of 25 and 42 kDa proteins arising from the proteolytic cleavage of the apo- enzyme separating the putative haem domain (25 kDa) and PQQ binding domain (42 kDa). Suitable isolation and purification protocols were developed to circumvent this problem. Partial protein sequencing has been performed from which DNA probes were deduced that allowed us to isolate a sequence in the gene. Preliminary attempts at expression in *E. coli* resulted in the formation of antibody detectable, inactive protein. Crystallisation of QH-EDH has been performed at the University of Groningen (Hol, Dijkstra, Huizinga). High resolution crystals have been obtained. Elucidation of the 3D-structure is expected in the near future. Preliminary experiments to establish the role of the haem and the PQQ-cofactors in the oxidation reaction suggest that PQQ is the primary site of alcohol oxidation while the haem is probably involved in the (internal) electron transport process. The mechanism of the consecutive oxidation of alcohol and aldehyde to the corresponding carboxylic acid end product has been elucidated. High enantioselectivity in the first step (enantiomeric ratio E = 100) is followed by low selectivity (E = 4) for the opposite enantiomer of the aldehyde. QH-EDH from *C. testosteroni* shows similarity to alcohol dehydrogenases from various acetic bacteria. Several of these bacteria are capable of enantioselective oxidation of 2,3-epoxypropanol, glycidol and solketal. Procedures for the production of homochiral glycidol and solketal have been patented. Andeno- DSM has investigated the possibilities for biotechnological applications.

Soluble quinoprotein glucose dehydrogenase from *Acinetobacter calcoaceticus* has been expressed in *E. coli*. In a parallel investigation the enzymatic properties have been determined. Test strips for glucose detection based on the use of this enzyme are presently developed by Boehringer Mannheim.

Molybdoprotein aldehyde dehydrogenase from *C. testosteroni* and from *Amicolatopsis methanolica* has been isolated, purified and characterised. Based on the rather limited substrate specificity, high enantioselectivity for aromatic (pro)chiral aldehydes was expected. So far, only low to insignificant selectivity could be determined for aldehydes of industrial interest.

A number of so called nicotinoprotein alcohol and aldehyde dehydrogenases, containing tightly bound NAD(P) as a cofactor, have recently been detected and characterised. Enzymatic activity could be demonstrated using *p*-nitroso-*N,N*-dimethylaniline, NDMA, as an electron acceptor. Methanol: NDMA oxidoreductase, MNO, from Gram-positive *A. methanolicus* was found to be able to oxidise methanol. In addition, MNO produces methylformate from methanol and formaldehyde, catalyses formaldehyde dismutation and NADH-dependent formaldehyde reduction. The bound cofactor is NADP. Another enzyme of this organism, NDMA alcohol dehydrogenase, NDMA-ADH, oxidises primary alcohols except methanol and contains NAD. The former enzyme appears to be highly enantioselective. A similar enzyme is probably involved in the enantioselective alcohol conversions for which *Rhodococcus erythopolis* has been patented.

The group in Lisboa has studied several enzymes from the groups in Delft and Munich. A detailed spectroscopic study was performed on enoate reductase (ER) from *C. tyrobutyricum*. UV/visible redox titrations of this enzyme were done to estimate the redox potentials of the cofactors. Analogue redox titrations were done using EPR as a monitoring technique. Two different signals were detected, one coming from a flavin, centred at g = 2.002 and another rhombic signal with resonances at g=2.013, 1.943, and 1.860. The change of the intensity of these signals with the redox potential has enabled us to determine the cofactors redox potentials. ^{57}Fe-enriched ER produced in Munich was characterised by Mössbauer spectroscopy. The study of its reduced and oxidised forms established that ER contains a [4Fe-4S] cluster, with unique spectroscopic properties. This has been interpreted as a 'non-all cysteine' coordination of the cluster. The spectral change in the presence of substrate suggests some relevant implications on the enzyme mechanism. An ENDOR study was performed on the reduced state of the enzyme. Good agreement was obtained between the 3D structure of a [4Fe-4S] centre and the spin populations, isotropic hyperfine couplings, initially assumed. ESEEM spectra were performed to probe nitrogen interactions with the cluster. Two nitrogens can be involved on the second sphere of coordination.

With carboxylic acid reductase from *C. thermoaceticum* preliminary EPR studies were performed. EPR signals attributable to tungsten and to $[4Fe\text{-}4S]^+$ clusters were observed. EPR studies on aldehyde dehydrogenase from *A. methanolica* were performed. Mo(V) signals of the slow and rapid type could be observed as well as two types of [2Fe-2S] centres with g max of 2.02 and 2.08. Alcohol dehydrogenase from *C. testosteroni* was characterised by UV/visible, EPR and NMR spectroscopies. Apo (without PQQ) and holoforms of the enzyme are clearly distinguishable by any of those techniques. The redox potentials of the heme is 60 mV more positive on the holoform. A PQQ titration followed by ^{1}H-NMR of the apoform shows that the binding of PQQ is very tight, apo and holoform are in slow chemical exchange ($k < 2 \times 10^2\ s^{-1}$).
A large modification in the chemical shift of one of the heme methyl groups (7 ppm) between the apo and holoforms is observed. All the spectroscopic techniques seem to indicate that the PQQ and the heme are in interaction or a big conformational change has occurred by the binding of the PQQ to the apoform. Ca, V and Co also seem to affect (not in such large extent) the chemical shift of the same methyl group.

At TNO the development of a QH-EDH based enzyme electrode was studied, the main subject being the development of a mediator free electrode for this quinoprotein. QH-EDH was immobilised on membrane electrodes coated with conducting polymers (polypyrrole). Adsorption of QH-EDH on the electrode is described by a Langmuir constant of 60 μg/ml and a maximum adsorption level of 26 $\mu g/cm^2$. No direct electron transfer between the enzyme and the functionalised electrode was observed. The immobilised enzyme is active when it is mediated by ferricyanide (Current density $\approx 15\ \mu A/cm^2$). Affinities of the enzyme electrode for various substrates are of the same order of magnitude as for the free enzyme. The interaction of redoxpolymers containing a redox active osmium-bipyridyl complex with a QH-EDH from the Delft group was studied. A functionalised polymer was constructed by coupling the osmium complex to a polyvinyl-pyridine backbone. QH-EDH was effectively wired to an electrode by combined deposition of an enzyme/polymer mixture on glassy carbon. The functionalised electrodes had initial current densities of $\approx 12\ \mu A/cm^2$. Typical half life times of these electrodes are about 3 hours. Bioelectrochemical activity of the electrode seems to be dependent on the amount of enzyme incorporated and the preparation scheme. Optimisation

of the electrodes focusing on the preparation scheme (temperature, pH) improved electrodes with respect to performance characteristics (initial current densities up to 40 $\mu A/cm^2$). Stability was improved by an additional crosslinking with glutaraldehyde. The half life time of the enzyme electrode was now 24-28 hours. Initial kinetics in an electrochemical cell resulted in K_m and V_{max} values comparable to the free enzyme. For 1-pentanol and ethanol K_m's for the enzyme electrode were 6μM and 820 μM respectively; for the free enzyme values 5μM and 1.5 mM, respectively.

ATO developed a simulation model for combined diffusion and enzyme catalysed processes which occur as a consequence of heterogeneous electron transfer reactions. It was validated with experimental results for the electron transfer of electrons from an electrode to viologen followed by reduction of hydrogenase and consecutive hydrogen production by the enzyme.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The regeneration of all 4 pyridine nucleotides has become possible by using crude extracts of *C. thermoaceticum*. Next year Boehringer Mannheim will offer AMAPOR activities from *C. thermoaceticum*. Also Boehringer develops test strips for the glucose detection by a quinoprotein glucose dehydrogenase. Establishing QH-EDH as excellent catalyst for the production of solketal and glycidol enantiomers. Discovery of new types of enzymes involved in alcohol/aldehyde conversions: molybdo and tungsten proteins and nicotinoproteins . Characterisation of various prosthetic groups in new enzymes by spectroscopic measurements. Construction of a mediatorless electrode for QH-EDH: direct communication between the enzyme and an electrode.

MAJOR COOPERATIVE LINKS

The group in Munich supplied the group in Portugal with pure enoate reductase and various samples of tungsten containing reversible aldehyde oxidoreductases, the Delft group with QH-EDH and others. Delft and Munich cooperated with respect to kinetic measurements. The group in Munich has sent several times samples of crude extracts of *C. thermoaceticum* to TNO. Group meetings were held in Munich, Delft, Braunschweig and Lisboa.

PATENTS

Enantioselective preparation of S-2R1,2R2-1,3-dioxolane-4-methanol and derivatives thereof. Inventor Duine et al.
Enzymatic, enantioselective preparation of R-glycidol. Inventor Duine et al.

PUBLICATIONS

Joint publications

Caldeira J., Feicht R., White H., Teixeira M., Moura J.J.G., Simon H., Moura I. *J. Inorg. Biochem.* (1993) **51**, 479
Somers W., van den Dool R.T.M., de Jong G.A.H., Jongejan J.A., Duine J.A., van de Lugt J.P. (1994), *Biotechnology Techniques,* **8**, 407-412
Caldeira, J., Feicht, R., White, H., Moura, J.J.G., Simon, H., Moura, I. (1992), *Towards Europe Second Italian-Portuguese-Spanish Meeting in Inorganic Chemistry*, Alfa Mar (Algarve) Portugal, 23 to 27 March 1992.
Caldeira, J., Feicht, R., White, H., Teixeira, M., Moura, J.J.G., Simon, H., Moura, I. (1993), Enoate reductase: *Workshop: Magnetic Spectroscopy on Bioinorganic Transition Metal Centers*, Homburg/Saar 24 to 27 February 1993.

Caldeira, J., Feicht, R., White, H., Teixeira, M., Moura, J.J.G., Simon, H., Moura, I. (1993), *9. Congresso Nacional de Bioquimica*, Alfa Mar (Algarve) Portugal, 5 to 8 December 1993.
Caldeira J., Moura J.J.G., Moura I., de Jons A., Devries S., Jongejan J.A., Duine A.J. NATO Advanced Research Workshop on Nuclear Magnetic Resonance of Paramagnetic Macromolecules, Portugal, June 4-8, 1994
Huber H., Caldeira J., Jongejan J.A., Simon H., *Arch. Microbiol.* (in press)

Individual publications

Bayer M., Günther H., Simon H. (1994a), *Appl. Microb. Biotechnol.* **42**, 40-45
Bayer, M., Schulz M., Günther H., Simon H. (1994b), *Appl. Microb. Biotechnol.* (in press)
Bystrykh L.V., Govorukhina N.I., van Ophem P.W., Hektor H.J., Dijkhuizen L., Duine J.A. (1993), *J. Gen. Microbiol.* **130**, 1979-1985
Bystrykh L.V., Vonck J., van Bruggen E.F.J., van Beumen J., Samyn B., Govorukhina N.I., Arfman N. Duine A.J., Dijkhuizen L. (1993) *J. Bacteriol.* **175**, 1814-1822
Duine J.A. (1991) Eur. *J. Biochem.*, **200**, 271-284
Duine J.A., Jongejan J.A., Geerlof A. (1994) *J. Pure Appl. Chem.*, **66**, 745- 752
Geerlof A., Jongejan J.A., van Dooren T.J.G.M., Raemakers-Franke n P.C., van den Tweel W.J.J., Duine J.A. (1994) *Enzyme Microbial Technol.*, in press
Geerlof A., van Stoorvogel J., Jongejan J.A., Leenen E.J.T.M., van Dooren T.J.G.M., van de Tweel W.J.J., Duine J.A. (1994), *Appl. Microbiol. Biotechnol.*, in press
Geerlof A., van Tol J.B.A., Jongejan J.A., Duine J.A. (1992) In: *Microbial Reagents in Organic Synthesis* (S. Servi ed.) NATO ASI Series C, Vol. 381, Kluwer Academic Publishers, Dordrecht pp. 412-420
Geerlof A., van Tol J.B.A., Jongejan J.A., Duine J.A. (1993), *J. Chromatogr.*, **648**, 119-129
Geerlof A., van Tol J.B.A., Jongejan J.A., Duine J.A. (1994), *Biosci. Biotech. Biochem., 58*, 1028-1036
Günther H., Simon H. (1994), *Biocatalysis* (accepted, invited review article on artificial mediators of redox enzymes)
Schulz M., Bayer M., White H., Günther H., Simon H. (1994a), *Biocatalysis* **10**, 25-26
van Ophem P.W., Bystrykh L., Duine J.A: (1992), *Eur. J. Biochem.*, **206**, 519-525
van Ophem P.W., Duine J.A. (1993). In: *Enzymology of carbonyl metabolism 4* (H. Weiner et al., eds.) Plenum Press, New York pp. 605-620
van Ophem P.W., van Beumen J., Duine J.A. (1993), *Eur. J. Biochem.*, **212**, 819- 826
Ye l:, Haemmerle M., Olsthoorn A.J.J., Schuhmann W., Schmidt H.-L., Duine J.A., Heller A. (1993), *Anal. Chem.*, **65**, 238-241

Enantioselective biotechnical resolution of racemic epoxides in the production of optically pure epoxides (BIOT CT-910269)

COORDINATOR:
J.A.M. de BONT, Agricult. Univ. Wageningen, NL
PARTICIPANTS:
Dr D.J. LEAK, Imp. College London, GB
M.M.R. da FONSECA, Inst. Sup. Técn. Lisbon, P
V.H.M. ELFERINK, Andeno B.V., Venlo, NL

BACKGROUND INFORMATION

A stereoselective enzyme has been discovered in a *Xanthobacter* species which resolves cheap racemic epoxide mixtures to very valuable optically pure epoxides to be used in subsequent industrial chemical synthesis of optically pure bio-active compounds. The objective of the project is to study epoxide-degrading enzymes with an emphasis on the enzyme from *Xanthobacter* Py2. The goal of the work is to arrive at an understanding on basis of which the industrial partner or eventually others can decide on industrial application of the method.

OBJECTIVES AND PRIMARY APPROACHES

Enantioselective epoxide degradation in *Xanthobacter* Py2 was studied both at the whole cell level, at the enzyme level, and at the level of its molecular genetics. Optimization of enzyme expression in whole cells was studied by the participant from Wageningen (chemostat cultures) but mainly by the Lisbon group. The Lisbon participant furthermore studied cell reutilization and immobilization for epoxide degradation purposes. Furthermore, the use of a second organic phase was studied there because epoxides are known to be toxic for whole cells. The Wageningen participant initially concentrated on the purification of the epoxide-degrading enzyme as well as on the unknown cofactor involved in the enzymatic epoxide dagradation. This work subsequently was carried out by the London participant. This group concentrated initially on the isolation of mutants of *Xanthobacter* Py2 for ^{13}C-NMR studies in the detection of transient ^{13}C-labelled intermediates. Considerable progress was made both in purifying the enzyme resulting in the determination of an N-terminal amino acid sequence of a protein and in identifying the unknown cofactor. Furthermore, a *Corynebacterium* was isolated which contained an epoxide hydrolase active against epoxides from cyclic compounds. The participant from Wageningen also isolated mutants for complementation work and mutants were exchanged between Wageningen and London. The complementation of mutants defective in epoxide degradation was succesful and the nucleotide sequence of the complementing DNA fragment was determined.

RESULTS AND DISCUSSION

Aspects 1-3 given below have been dealt with at Imperial College in London. Aspects 4 and 5 have been investigated at the Wageningen Agricultural University and aspects 6-8 have been studied at the Instituto Superior Técnico in Lisbon.

(1) Isolation of mutants of *Xanthobacter* Py2 for ^{13}C-NMR studies

At the outset of this project the steps in the metabolism of propene oxide had not been established. By following the metabolism of a ^{13}C-labelled substrate it is sometimes possible to detect transient ^{13}C-labelled intermediates. However

because of the relative insensitivity of ^{13}C-NMR it is often better to allow intermediates to accumulate in pathway mutants. For this reason we isolated a series of mutants which had lost the ability to grow on propene oxide. Some of these were completely unable to breakdown propene oxide while a few degraded the epoxide slowly. The latter are interesting candidates for ^{13}C-NMR work. However, because of the progress of the Wageningen group in establishing a potential route for degradation of the epoxide, we have delayed undertaking the ^{13}C-NMR until the protein purification work was complete.

(2) Studies on the epoxide degrading enzyme from *Xanthobacter* Py2

Preliminary work had demonstrated that enzyme activity required a low molecular weight unidentified soluble cofactor (LMF) as well as the protein components (HMF). Studies in Wageningen and London demonstrated that the enzyme could be inhibited by the sulphydryl reagent N-ethylmaleimide and that dithiothreitol (DTT) and other reduced sulphydryl compounds could replace the LMF in the assay. However the LMF did not appear to be sensitive to sulphydryl binding reagents.

The ability of DTT at low concentrations (high concentrations attack the epoxide chemically) to replace the LMF considerably facilitated the purification of the enzyme. This has now largely been achieved using the protocol shown on the next page.
From the conditions required for purification, fraction B appears to be very hydrophobic. Native and SDS-PAGE demonstrated that it was >95% pure after the HIC step and was comprised of a monomer of approximately 62 kilodaltons. Spectrophotometric analysis revealed that fraction B had a chromophore typical of a flavoprotein. An N-terminal amino acid sequence has been obtained by automated Edman degradation, vis:

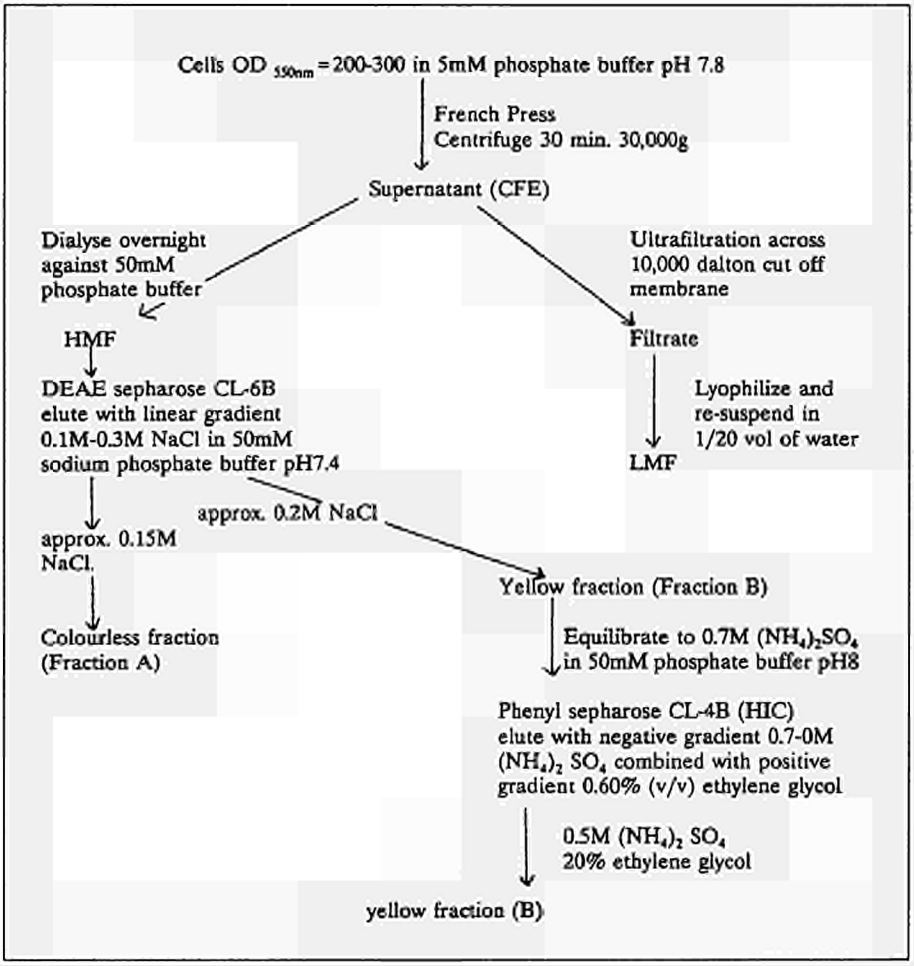
M-K-V (or M)-C(?)-N-A-R-N-D-H--L-L-I(or L)-x-D

Fraction A arising from the initial ion exchange was colourless and, assuming that it was the major band appearing on native and SDS-PAGE was >80% pure, probably a homodimer/trimer comprised of 44 kilodalton subunits. Despite trying a number of techniques including hydrophobic interaction, ion exchange (on a different matrix) and hydroxylapatite chromotography it has not been possible to remove the 'contaminating bands'.

The LMF is available in very small amounts, one 15L fermentation yielding sufficient for about 60 assays. As the Wageningen group tried a variety of recognised cofactors without success, we have started to purify the LMF for structural analysis. Preliminary results from solid phase extraction (SPE) look promising suggesting a bulk ion exchange step followed by reverse phase HPLC. The latter should enable direct interfacing to the LC-MS facility available in this department for preliminary structural analysis.

(3) Isolation of a cyclohexene oxide degrading *Corynebacterium* sp.

Epoxide degradation via hydrolysis is well characterised in eukaryotes but there are few reports of bacterial epoxide hydrolases. Initial attempts to isolate bacteria degrading cyclic and linear subterminal epoxides revealed that unless strict attention was paid to the problems of substrate toxicity and the propensity for epoxides to spontaneously hydrolyse in aqueous media, little progress would be made. However we devised a two liquid phase isolation system in which the epoxide dissolved in an organic phase in a central well of an erlenmeyer flask, partitioned via the gas phase into the growth medium, achieving a steady supply of epoxide at



non-toxic levels. Using this system we isolated a cyclohexene oxide degrading *Corynebacterium* sp. which contains an epoxide hydrolase, active against a wide range of cyclic and internal epoxides.

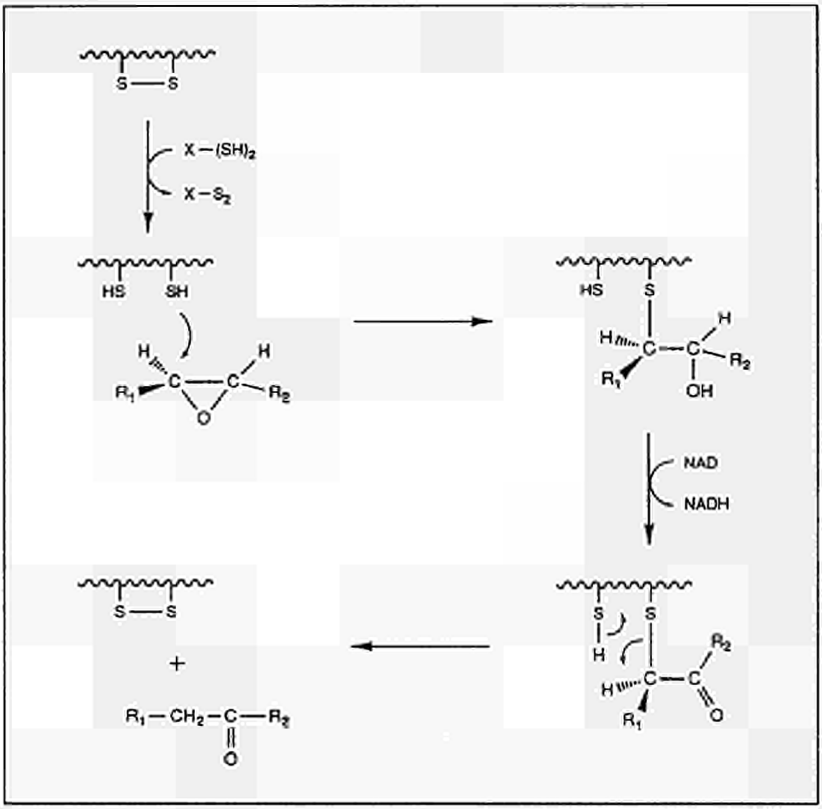
(4) Degradation of epoxides in cell extracts from *Xanthobacter* Py2.

A broad range of epoxyalkanes was converted into the corresponding ketones by cell extracts of *Xanthobacter* Py2. Both 1,2-and 2,3-epoxyalkanes were degraded and in addition, the degradation of 2,3-epoxyalkanes in all cases was highly enantioselective.

Conversion of a deuterium-labelled substrate indicated that the ketone product was probably formed indirectly via an hydroxy intermediate.

Degradation of epoxyalkanes by *Xanthobacter* Py2 was dependent on both NAD and another, not yet identified cofactor which was present in the low molecular weight fraction (LMF) of propene-grown cells. It is proposed that the LMF was

involved in a reductive reaction step since it could be replaced by dithiothreitol (DTT) and various other dithiol compounds. Epoxyalkane-degrading activity was inhibited by the sulfhydryl blocking reagent N-ethylmaleimide (NEM). Inhibition by NEM and stimulation by LMF, DTT and other dithiols was only effective in the simultaneous presence of NAD.



Based on these results, a mechanism for the enzymatic degradation of epoxides was proposed (see scheme). The unknown cofactor is represented by $X\text{-}S_2/X\text{-}(SH)_2$.

(5) Complementation of *Xanthobacter* Py2 mutants defective in epoxide degradation

Three *Xanthobacter* Py2 mutants designated M3, M8 and M10 lacking epoxyalkane-degrading activity were isolated and characterized. All mutants were able to grow on ketones, the degradation product of epoxyalkane conversions. Furthermore, they contained the active form of the low molecular weight fraction (LMF) necessary for activity. The mutant *Xanthobacter* Py2M10 was used for screening of a gene bank. Three cosmids were found complementing the mutation in M10. These cosmids also complemented the mutation in the mutant M8, but not in mutant M3. Activities in crude extracts of 1,2-epoxypropane grown complemen-

ted mutants were similar to the wild type activities. Surprisingly, M10 transformed with complementing cosmid pEP9 showed a constitutively expressed epoxyalkane-degrading activity, which is not observed in the wild type Py2. The cosmid pEP9 was conjugated to *Xanthobacter autotrophicus* GJ10 which is not able to degrade epoxyalkanes. In crude extracts of *Xanthobacter autotrophicus* GJ10 harbouring pEP9 epoxyalkane-degrading activity was demonstrated, however, only after the addition of the LMF from *Xanthobacter* Py2. Hybridisation experiments demonstrated an overlap present on complementing cosmids pEP1, pEP3 and pEP9. Subcloning revealed a 4.8 kb *Eco*RI-*Hind*III fragment necessary for complementing the mutant M10. In the sequence of this fragment four different open reading frames were found.

The largest open reading frame showed significant homology with lipoamide reductase.

(6) Growth of *Xanthobacter* Py2 on 1-alkenes

The propene isolated strain *Xanthobacter* Py2 was able to grow on the following 1-alkenes: C5, C6, C8, C10, C12, C14 and C16. The experimental biomass yield for growth under alkene limiting conditions could only be calculated on alkenes with a low boiling point, so that substrate consumption could be measured by head-space analysis. The results in Ceq/Ceq were as follows: propene -0.53; 1-pentene — 0.51; 1-hexene — 0.47.

In higher alkenes (C10-C16) cells clustered together in long chains or small granules during the initial period of the exponential phase. Further on the aggregation became more severe, keeping the large majority of cells in a few clusters.

We tried to prevent aggregation by adding an inert organic phase to the medium, namely n-dodecane. Still some aggregates were formed, although smaller than those in hexadecene alone. However, analysis of the organic phase revealed that no hexadecene had been consumed. This suggested that dodecane had been used for growth. Further experiments showed that cells adapted to either 1-pentene or 1-hexadecene grew well on aliphatic alkanes with chain lengths ranging from C6 to C16. No growth occurred on cyclohexane, benzene, toluene and xylene.

(7) Degradation of 1,2-epoxyalkanes by alkene grown *Xanthobacter* Py2

To study the influence of the chain length of the 1-alkenes, used as a growth substrate, on the degradation capacity towards 1,2-epoxyalkanes, *Xanthobacter* Py2 cells grown on 1-alkenes ranging in chain length from C3 to C16 were used in degradation reactions (Table 1). Surprisingly, cells grown on alkanes (C6 — C16) also showed degrading activity towards 1,2-epoxides. Degradation rates for 1,2-epoxyhexane ranged from 2 to 5 nmol/(min * mg protein).

(8) Cell reutilization and immobilization for epoxide degradation purposes

Attempts have been made to immobilize *Xanthobacter* Py2 cells in gels (using cells harvested at the beginning of the stationary phase) and also *in situ* in polyurethane reticulated cuboids of different sizes and pore diameters. The most promising results were obtained with *in situ* immobilization in which case immobilized cells showed a specific degrading activity comparable to that of freely grown cells.

The purpose of subsequent research was to understand the reasons underlying the poor operational stability of cells. The studies were carried out with suspended cells.

Table 1: Degradation rates of 1,2-epoxyalkanes by alkene grown *Xanthobacter* Py2. Units in nmol/(min*mg protein)
Chain length of growth substrate (1-alkene)

Substrate chain length (1,2-epoxyalkane)	*C3*	*C5*	*C6*	*C8*	*C10*	*C12*	*C14 orC16*
C3	68	95	60	5	2	<0.5	ND
C4	49	55	44	4	1	<0.5	ND
C6	15	22	17	3	2	5	ND
C8	4	6	6	4	3	9	ND
C10	0	<0.5	<0.5	1	1	4	4-8
C12	0	0	<0.5	<0.5	<0.5	1	3-6
C14	0	0	0	0	0	1	ND

No correlation was found between the viability of cells and their degrading activity.

The results obtained suggest that the enzyme(s) responsible for the epoxide degradation
(i) is (are) irreversibly inhibited or
(ii) need a co-factor which resting cells are unable to regenerate.

Apparently, new enzymes and/or co-factors have to be primarily synthesized.

Remarkable is that resting cells are unable to maintain their epoxide degrading activity in the presence of epoxides, even though the latter are intermediate metabolites during growth.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The observation that dithiothreitol replaces the unidentified soluble cofactor has allowed the purification and characterization of the proteins involved in epoxide degradation. The complemenation of mutants defective in epoxide degradation was succesful and the nucleotide sequence of the complementing DNA fragment was determined.

In view of industrial applications enantioselective degradation of epoxides by whole cells was further optimized.

MAJOR COOPERATIVE LINKS

Strong cooperative links between participants were maintained during the full period of the project and five joint meetings were held.

JOINT PUBLICATIONS

J. Swaving, C.A.G.M. Weijers, C.K. Chan Kwo Chion, D.J. Leak, A.J.J. van Ooyen and J.A.M. de Bont. Initial characterization of the enzyme and cloning of genes involved in the enantioselective epoxyalkane degradation by *Xanthobacter* Py2. Biocatalysis, 1994, **10**, 227-232.

INDIVIDUAL PUBLICATIONS

F. van Keulen and M.M.R. da Fonseca. Batch cultivation of *Xanthobacter* Py2 on 1-pentene: growth yields and cell activity for the degradation of 1,2-epoxyhexane. Submitted to Biotechnology Letters, 1994, **16**, 989-994.

C.A.G.M. Weijers, H. Jongejan, M.C.R. Franssen, A. de Groot and J.A.M. de Bont. Dithiol- and NAD-dependent degradation of epoxyalkanes by *Xanthobacter* Py2. Applied Microbiology and Biotechnology, (in press).

J. Swaving, C.A.G.M. Weijers, A.J.J. van Ooyen and J.A.M. de Bont. Complementation of *Xanthobacter* Py2 mutants defective in epoxyalkane degradation; expression and nucleotide sequence of the complementing DNA fragment. Microbiology, (in press).

Glycosyltransferases from streptomycetes as tools in biotransformations (BIOT CT-900155)

COORDINATOR:
W. PIEPERSBERG, BUGHW FB9, Wuppertal, D
PARTICIPANTS:
A. JIMÉNEZ, UAM and CSIC Madrid, E
E. CUNDLIFFE, Univ. Leicester, GB
S. GRABLEY, B. BRÄU, R. MARQUARDT, Hoechst AG, D

BACKGROUND INFORMATION

Glycosylation of low-molecular weight bioactive molecules, especially by rare sugars, is frequently an essential mean to determine their (e.g.) pharmacological specificity and pattern of action. More than 80 different 6-deoxyhexose (= 6DOH) derivatives occur in secondary metabolites from actinomycetes and other taxa. Among the sugar moieties found in streptomycete antibiotics these form the largest and most varied group. Therefore, the 6DOH pathways are of high interest for biotransformations and pathway engineering.

OBJECTIVES AND PRIMARY APPROACHES

The aims of the current study were as follows

(1) Identification, cloning and expression of the (dTDP-)6DOH biosynthetic genes and enzymic analysis of the glycosylation enzymes — including 6DOH-activating, -modifying and -transfering (glycosyltransferases = GTs) enzymes — from various strains of antibiotic-producing streptomycetes;

(2) preparation of GT-substrates (aglycones) and co-substrates (activated 6DOH, derivatives of dTDP-D-glucose) at analytical and preparative scale;

(3) further enzyme characterization via testing GT specificity for aglycones and/or 6DOHs requiring GT-overexpression and purification.

RESULTS AND DISCUSSION

1. Analysis of 6DOH genes from (3'-hydroxy-)streptomycin producers, design of probes and screening for actinomycete 6DOH genes, and preparation of dTDP-6DOH (BUGH)

The work of the BUGH group has focussed in the past 3 years on the producers of streptomycin-related antibiotics among the streptomycetes, mainly *S. griseus* N2-3-11 (streptomycin), and *S. glaucencens* GLA.0 (5'-hydroxystreptomycin). In the respective *str/sts*-gene clusters 30 different genes have been identified so far. They contain, besides those known to be involved in dTDP-6DOH metabolism and widely spread among other bacteria (Fig.1,2; Tab.1), other related genes probably involved in the NDP-N-methyl-L-glucosamine pathway (e.g. *strPQ, strX*). With many relationships to the gene clusters for extracellular polysaccharide synthesis (e.g. the *rfb*-genes for LPS O-chain biosynthesis) of Gram-negative bacteria these encode a set of enzyme families each engaged in very similar sugar (or cylitol) conversions (Tab.1). This is especially exemplified by the families of gene products involved in dTDP-6DOH biosynthesis, such as those related to StrD, StrE, StrM. Also, a new group of putative secondary metabolic aminotransferases (SMATs; e.g.

StrS, Prg1, TylB; cf. CSIC and ULEI) has been detected, which besides other sugar or cyclitol derivatives seem to use dTDP-6DOH ketone intermediates as their typical substrates. However, no GT could be identified by sequence similarity (e.g. related to the putative GTs identified by ULEI and HOE) among the 30 Str (Sts) gene products encoded in the DNA of *S. griseus* and *S. glaucencens*. Therefore, to

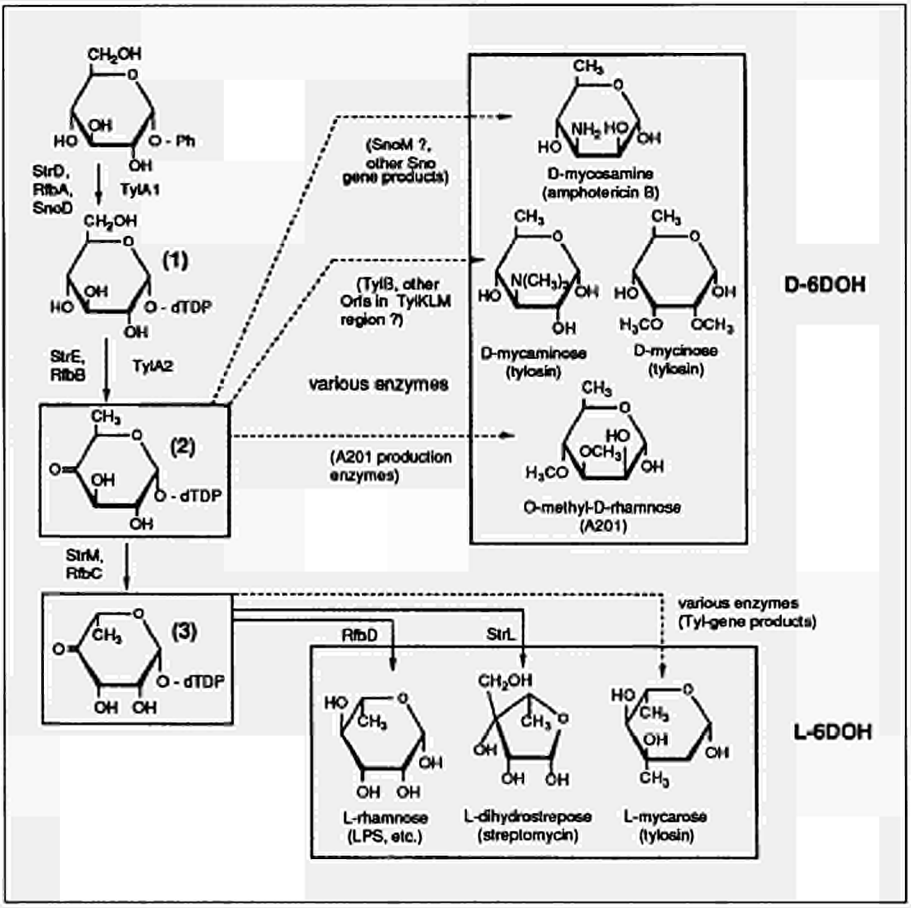


Fig. 1. 6DOH-pathways for relevant GT substrates. Identified biosynthetic enzymes are indicated by the symbols of the respective gene products. The crucial intermetiates dTDP-D-glucose (1), dTDP-4-keto-6-deoxy-D-glucose (2), and dTDP-4-keto-L-rhamnose (3) have been prepared at preparative scale.

enable GT assays — e.g. to test for the hypothetical function of the StrH protein as the dTDP-L-dihydrostreptosyl GT — at the enzymatic level major efforts at BUGH were focussed on the production of the dTDP-6DOH intermediates 1 to 3, as well as dTDP-L-dihydrostreptose and dTDP-L-rhamnose (Fig. 1). This was achieved recently by overexpressing in E. coli or *S. lividans* of the enzymes StrDEML (from *S. griseus*) and RfbABCD (from Salmonellae), respectively (Fig.1,2; Tab.1). For the *rfb*, but not for the *str* genes, stable expression clones giving very high enzyme yields could

be contructed in T7 polymerase/promoter expression systems. StrE and StrL were expressed in *E. coli* at a much lower scale; StrE could also be expressed in *S. lividans* in detectable amounts. It could be demonstrated by various biochemical analyses that our original assignment of enzyme functions was correct, especially for the StrM/RfbC and StrL/RfbD pairs of related proteins (see Fig.1,2; Tab.1) which were controversially assigned in the literature.

Tab. 1. Protein families involved in NDP-sugar and cylitol metabolism [a]

Family (type protein)	*general function*	*gene products dTDP-6DOH metabolism (examples)*	*examples from other (NDP-sugar or cyclitol) pathways*
StrD-family	NDP-sugar synthases (pyro-phosphorylases	StrD, TylA1, SnoD, RfbA, (LmbO?)	StrQ, RfbE, GlgC)
StrE-/GalE-family	NDP-hexose 4-(2-)oxido reductases (epimerases or dehydrogenases	StrE, TylA2, RfbB, SnoE, StrL, RfbD, (LmbM?)	StrP, GalE, RfbJ, Rfb
StrM-family	NDP-4-keto-hexose isomerases (epimerases)	StrM, RfbC, (Tyl-Orf6*?), (SnoM?)	StrX
StrS-/EryC1-family	(NDP-)sugar or cyclitol aminotransferases (putative) or dehydrases	(TylB?), (EryC1?), (DnrJ?), (LmbS?)	AscC, RfbH, (StrS, StsA, StsC?) (Prg1?)
MgtA-family	NDP-hexose glycosyl-transferases	(Tyl-Orf2*?), (SnoT?)	MgtA, flavonol-GT, various sterol-GTs

[a] Gene products form other pathways as those mentioned in the text are: Lmb = lincomycin, Glg = glycogen; Gal = galactose; Ery = erythromycin; Dnr = daunorubicin

2. Identification and cloning of A201A production genes (CSIC)

The CSIC group screened for the 6DOH (O-methyl-D-rhamnose; Fig.1) genes/enzymes involved in biosynthesis of the nucleoside antibiotic A201A in *S. capreolus*. These were found in a total of about 48 kb of non-contiguous DNA from *S. capreolus* by (a) cloning in cosmid pJAR4 of the genetic determinat(s) for self-resistance to A201A from *S. capreolus* and (b) hybridization with DNA probes from the streptomycin (from BUGH) and puromycin biosynthetic gene clusters of *S. griseus* and *S. alboniger*, respectively. Two groups of A201A resistance-mediating cosmid clones, each represented by two independent isolates (group 1: pCAR13/14; group 2: pCAR11/23), were obtained. Subcloning and sequencing of a 4 kb fragment from pCAR14 localized an A201A resistance determinat (*ard1*, encoding a 508 aa polypeptide) which belonged to the ABC-transporter superfamily. Immediately 5' to *ard1* another gene was detected with a deduced aa sequence highly similar to the Pur3 protein from *S. alboniger*, a member of the inositol monophosphatase family, which is thought to be involved in the biosynthesis of the 3'-amino-3'-deoxyadenosine moieties of puromycin and A201 in both *S. alboniger*

and *S. capreolus*, respectively. This indicated that the DNA in group 1 cosmids carry genes of the A201A biosynthetic cluster. Further hybridization experiments on total genomic DNA from *S. capreolus* and group 1 and 2 cosmids with gene-specific probes from the *strDELM, strH* and *strS* genes from *S. griseus* and the *pur3, 5, 6, 10* and *prg1* genes from *S. alboniger* detected hybridizing bands with *strH* and *strM* in group 2 and 1 cosmids, respectively, which were cloned each on 1.6 kb fragments and are presently analysed. Also, in total DNA hybridizing bands were detected with all the other probes tested, except for *strD*. In conclusion, a group of genes belonging to the biosynthetic gene cluster for A201A from *S. capreolus* has been cloned.

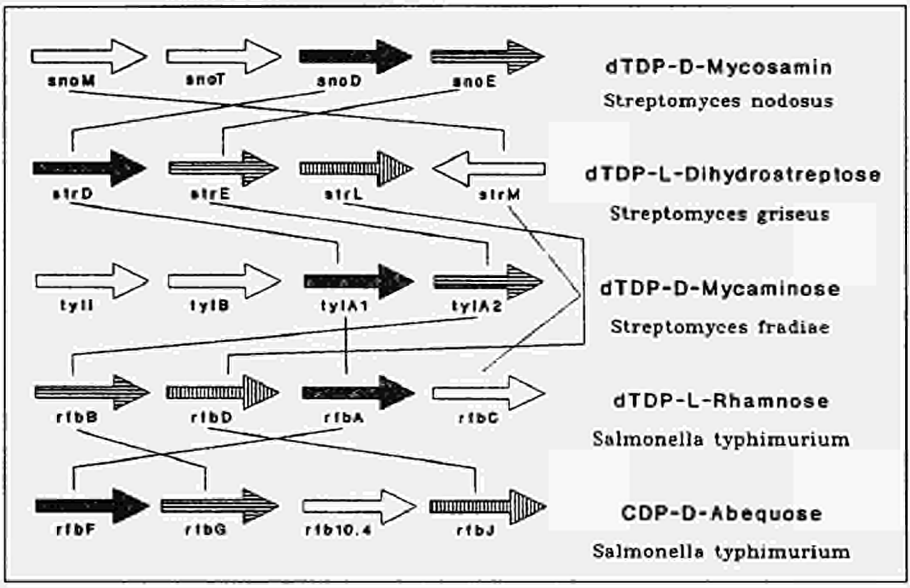


Fig. 2. Arrangement of the basic dTDP-(CDP-)6DOH production genes from various organisms and relationships among them. The gene order in the respective clusters is given schematically. For details see text.

3. Analysis of 6DOH production and GT genes from the tylosin producer *S. fradiae*; characterization of a macrolide glucosyltransferase (MgtA) from *S. lividans* (ULEI)

Considerable sequence analysis has been carried out within the tylosin biosynthetic gene cluster of *S. fradiae*, aided by the gene probes from BUGH, in order to detect the 6DOH genes (Fig.1). The *strD* and *strE* probes detcted the *tylIBA* region in which 5 genes (named *tylIBA1A2*-ORF5) were analysed (Fig.2; Tab.1). within and, latterly, 6 additional genes (ORFs 1*-6*) within the *tylLM* region that is located some distance from *tylIBA* on the other side of the polyketide synthase gene cluster (cf. public. no. 6,7). ORF1 *(tylI)* encodes a putative C20-hydroxylase of tylactone, ORF2 *(tylB)* belongs to the hypothetical SMAT family (Tab.1) and could be the aminotransferase in the dTDP-D-mycaminose pathway (Fig.1); ORF3 (tylA1) and ORF4 *(tylA2)* encode the *StrD* and *StrE*-related dTDP-glucose synthesising and converting enzymes (Fig.1, Tab.1). For *TylA1* this was confirmed by expression in *E. coli* and assay of its function. ORF5 encodes a thioesterase that

may be involved in release /cyclisation of the completed aglycone during tylactone formation. The 6 putative genes sequenced in the *tylLM* region are currently attributed the following functions: ORF1* (no function known), ORF2* (GT; probable mycaminosyl transferase; cf. Tab.1), ORF3* (no function known; sequence similarity to a spiramycin biosynthesis gene), ORF4* (oxidoreductase; possible ring oxidase of tylactone at C20 creating an aldehyde group), ORF5* (possible N-methyl transferase in the formation of the dTDP-mycaminose; ORF6* (*tylK*; possible 3'5'-epimerase in mycarose biosynthesis) (Fig.1). — Also, a resistance gene *(mgtA)* from *S. lividans*, was shown to encode a macrolide 2'-O-glucosyltransferase (MGT) enzyme that inactivates macrolide antibiotics by transferring glucose onto the mycaminosyl or desosaminyl residues of the drug molecules. Since this was the first GT gene to be positively identified among the actinomycetes it was tested as a DNA probe to identify GT genes in antibiotic biosynthesis clusters. Though the *mgtA* gene found single hybridisation targets in the genomes of several other streptomycetes, no hybridisation target was found in DNA from *S. fradiae* (organism studied at ULEI) nor in the *str* gene cluster of streptomycin producers (together with BUGH). However, a gene located with the *tylLM* region of the *S. fradiae* genome (ORF 2*) has revealed extensive sequence similarity to *mgtA* and *snoT* (see HOE) leading us to the provisional hypothesis that ORF 2* encodes the GT that adds the 6DOH mycaminose (Fig.1) to the emerging tylosin molecule. In order to test that hypothesis, it will be necessary to prepare the (presently unavailable) co-substrate dTDP-mycaminose for use in enzyme assays with the purified product of ORF 2*. Such studies are in prospect in view of the now available dTDP-6DOH precursors/intermediates (cf. BUGH).

4. Amphotericin B: Identification of D-mycosamine production genes *(sno)* and mutant induction in *S. nodosus* DSM40109, preparative studies on the aglycone (HOE)

Work at HOE aimed at the isolation of the D-mycosaminyl-GT from amphotericin B producers for its use as a tool for biotransformations of other (polyene) aglycones. HOE in collaboration with BUGH identified the 6DOH production genes *snoD* and *snoE* in strain *S. nodosus* via their hybridization to the respective *strD* and *strE* genes. The *snoD* gene and two adjacent genes, *snoM* and *snoT*, were cloned partially and analysed by sequencing on a 2.6 kb *Bam*HI fragment (cf. Fig. 2). Interestingly, the product of the *snoD*-adjacent gene, *snoT*, had significant sequence similarity to Mgt, Tyl-Orf2* (see section 3) and more distantly to various other GTs from plants, viruses, and animals, especially in a C-terminal motif with the consensus sequence:
W-PQ--VL-H----AFVTH-G--S---GL---VPM---P--GDQ (dashes = not highly conserved residues).
This suggested SnoT to be the D-mycosaminyl-GT whenever the isolated DNA-segment belonged to the amphotericin B production cluster. To prove this, attempts to construct disruption mutants in the *snoD* gene were carried out via use of a temperature-sensitive suicidal vector containing truncated forms of the gene. Transformants containing such plasmids mostly turned out to be pleiotropically altered, segregating various phenotypic variants after treatment at non-permissive temperature. Some derivatives of this kind had still retained the recombinant plasmid. Others had lost the amphotericin B production which did not correlate with the plasmid content. One clone produced a novel, amphotericin B-related secondary metabolite, but no strain could be identified which produced only the aglycone amphoteronolide B.

MAJOR SCIENTIFIC BREAKTHROUGHS

Three putative GT-encoding genes could be identified in streptomycets. A large variety of genes/enzymes for the synthesis of various dTDP-6DOH derivatives were cloned and analysed. Some dTDP-6DOH co-substrates for glycosylation reactions have been prepared quantitatively.

MAJOR COOPERATIVE LINKS

Biological (bacterial strains, substrates), genetic (DNA-probes, oligonucleotides), and analytical material and methods, as well as resulting data were exchanged continously among all participants. Regular meetings of 1 to 2 days were organized twice per year, and took place in all three participating European countries. The project was presented at a workshop organized for all the biotransformation projects at Braunschweig (FRG) in 1993.

PUBLICATIONS AND PATENTS

Joint publication

Stockmann, M., et al. 1991. German and international patent application. 41 30 967 (HOE 91/F 300).

Individual publications

Stockmann, M., W. Piepersberg 1992. FEMS Microbiol. Lett. **90**: 185-190.
Retzlaff, L., et al. 1993. pp. 183-194 in: Hegeman, G.D., R.H. Baltz, and P.L. Skatrud (Eds.) Genetics and Molecular Biology of Industrial Microorganisms. Proceedings of the 5th ASM Conference. ASM, Washington DC.
Piepersberg, W. 1994. Pathway engineering in actinomycetes. CRC Crit. Rev. Biotechnol. **14**: 251-285.
Cundliffe, E. 1992. Antimicrob. Agents Chemother. **36**: 348-352.
Cundliffe, E., L.A. Merson-Davies, and G.H. Kelemen. 1993. pp. 235-243 in: Hegeman, G.D., R.H. Baltz, and P.L. Skatrud (Eds.) Genetics and Molecular Biology of Industrial Microorganisms. Proceedings of the 5th ASM Conference. ASM, Washington DC.
Cundliffe, E., et al. 1994. Mol. Microbiol. (in press).
Lacalle, R., et al. 1992. EMBO J. **11**: 785-792.
Tercero, J.A., et al. 1992. FEMS Microbiol. Lett. **90**: 203-206.

AREA B:

ENABLING TECHNOLOGIES

DNA SEQUENCING

Alternative methods of DNA-sequencing (BIOT CT-900252)

COORDINATOR:
R. CARR, CAMR, Wiltshire, UK
PARTICIPANTS:
K. HOLSTROM, Biotechnology Institute, Lyngby, DK
G. CHARALAMBOS, Optimum Ltd., Athens, GR

BACKGROUND INFORMATION

To facilitate the analysis of DNA a rate capable of supporting our rapidly growing understanding of molecular biology and genome research, much of effort has been aimed at improving methods of sizing, mapping and sequencing of DNA fragments. While the rate and DNA size range over which conventional gel electrophoresis based methods of DNA analysis can be carried out has dramatically increased over recent years (as exemplified by the Généthon approach to the Human Genome Project), there is still a great deal of interest in the development of techniques by which DNA molecules can be analysed on an individual basis. A number of methods have been tried including Scanning Tunnelling and Scanning Force Microscopy, Flow Cytometry and video fluorescence microscopy all of which allow (to varying degrees) the structure of single DNA molecules to begin to be determined. The ability to resolve the internal structure (sequence) of the molecule is, as yet, far in the future but indications are that such approaches will become feasible.

OBJECTIVES AND PRIMARY APPROACHES

The initial objective of this project was to develop a solution phase method of analysing (e.g. sequencing) DNA. The approach adopted was to employ electron beam fabricated optical apertures of sub-wavelength dimensions sufficiently small to allow the molecule to be interrogated directly and, through employing base specific labels, allow the internal structure (e.g. sequence) to be determined. A change in participant structure (withdrawal of sub-contractor) necessitated an alternative approach to fulfilling some of the original project aims. Accordingly, initial work on the development of a non-laser based gel scanner was replaced with a comprehensive feasibility study on the use of the CAMR optical aperture detection system as a method to enable rapid, one-step **mapping** of DNA (e.g. large (10sKb) DNA fragments).

While CAMR developed the nanofabrication techniques necessary to fabricate arrays of submicron apertures and to identify alternative sources of suitable pore containing membranes through which DNA molecules could be electrophoresed, the Biotechnology Institute, Copenhagen, worked on methods by which suitable fluorescent labels could be attached to the DNA molecules. Originally looking a single base specific labels they also worked on the development of fluorescent labels to be attached to PNA oligo probes for mapping purposes which became another project aim during the second year.

The development of software and hardware capable of allowing the optical signals generated by the aperture array system was the responsibility of Optimum Ltd., of Athens. They undertook to design the data analysis programmes by which DNA samples could be analysed in the optical system in use at CAMR.

RESULTS AND DISCUSSION

A. CAMR (UK) has been working on the following areas:

1. Device Fabrication.

A variety of methods by which membranes and films containing sub-micron apertures could be fabricated or obtained was investigated. High energy electron beam writing was used to fabricate 50nm pores in thin (200nm) films of Al2O3 but subsequent illumination of the surface failed to detect the apertures in sufficient detail. (Devices fabricated by eximer laser photoablation of polyimide sheet and containing 5x5 arrays of 5μm pores were also assessed and despite proving suitable, supplies were unreliable. An alternative membrane type (polycarbonate track etched membrane) that is both commercially available and inexpensive was identified and proved adequate for our initial studies. These membranes, moreover, can be supplied with a wide range of sub-micron pore diameters, coated with a 200nm gold film or supplied pre-stained black for optical opacity and have excellent insulating properties (which will prove important in electrophoresis studies). Devices (100+) comprising metallised (60nm Cr) quartz substrates in which arrays of apertures of varying sizes have been fabricated using electron beam lithography and wet and plasma etching, have been prepared at CAMR. These devices were used to assess the detectability of sub-micron particulates by the detector system described below.

2. Construction of optical illumination and detection system.

Both back-illumination and grazing incidence illumination have been considered. For model systems in which metallised optical substrates bearing sub-micron apertures were used to determine optical sensitivity to fluorophores the use of low powered HeNe and Ar beams at grazing incidence have proved most sensitive. Employing chrome as the metal gives the advantage of durability and flexibility in accuracy of launch angle. Oblique back-illumination with low powered Ar is likely to be employed when track etched membranes are used in fluorescent studies.

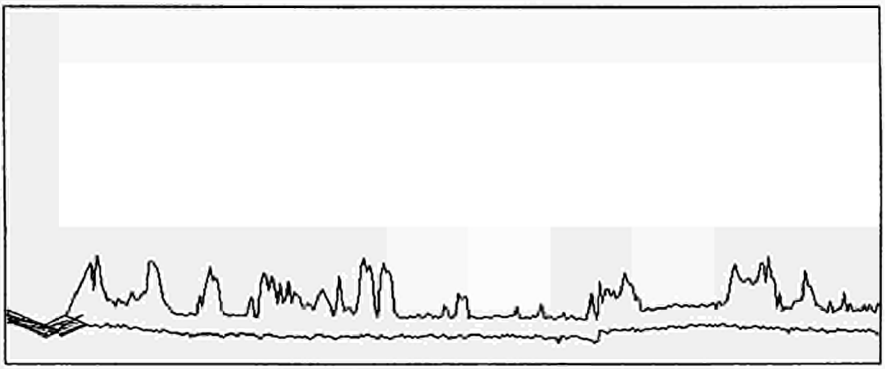
Detector systems based on avalanche photodiodes (APD) and charge coupled devices (CCD) have both been designed and the latter is being used to follow intensity fluctuations in light emanating from sub-micron apertures in response to the presence of labelled and unlabelled sub-micron particles diffusing in their immediate vicinity. Output from the CCD detector array is monitored by either video line selector which allows the intensity of light from any given aperture to be monitored in real time or by a special data acquisition system (Brian Reece Scientific Ltd, UK) allowing digital processing of the image (e.g. averaging, expansion, background suppression etc..). This latter system also allows monitoring of individual apertures in real time though limited by the frame refresh rate of the CCD array (i.e. 50Hz). The CCD detection system has proved successful in our initial work and the image analysis system is currently being modified in software (Optimum Ltd, Greece) to enable data collection and storage to be improved.

3. Model particulates detection.

A variety of model sub-micron particulates have been studied in order to define the sensitivity of the detection system and characterise the optical response of the apertures:

(i) Calibration polystyrene microspheres.
(ii) Au sol.
(iii) Fluorescently labelled microspheres.

Polyimide membrane systems, back illuminated with a white light source, were assessed using unlabelled 2m polystyrene microspheres prior to using polycarbonate track etched membrane containing pores of 200nm diameter. These membranes were stained with Ergolan black to enhance contrast and proved successful in detecting 400nm microspheres diffusing in aqueous conditions. Current work involves using similar membranes coated with 200nm Au and illuminated by grazing incidence Ar laser in order to achieve higher resolution. The figure 1 below shows intensity fluctuations arising from the scattering of light from 300Å gold sol particles.



This level of signal is remarkable in that these particles are visible using only a conventional optical microscope (x40 objective) mounted with a simple CCD camera. Moreover, the illumination intensity is mW only, suggesting a future instrument based on these illumination principles could be robust and low cost. The exact mechanism by which such low levels of incident illumination and modest magnifications can visualise such small particles is currently under investigation.

Fluorescently labelled microspheres (410nm labelled with fluorescein) have been monitored using grazing incidence illumination (fibre-optic mediated 488nm Ar ion beam) of an Al coated glass substrate containing apertures of 200nm diameter but photobleaching of the fluorophore in this method proved too rapid for quantitative data to be obtained.

Samples of 48.5kb lambda DNA were stained with the newly developed intercalating DNA dyes TOTO and YOYO (Molecular Probes Inc.) and are being used to determine the correlation between fluorescence signal with DNA fragment size. Significantly, recent work by Keller and his coworkers [Nucl Acid Res. (1993), 21(4), 803-806] have demonstrated that such a staining protocol can be used to accurately and rapidly size individual lambda DNA fragments over the 10-50kb size range with better than 2% accuracy using a sophisticated flow cytometer proving the feasibility of such an approach. CAMR efforts will however, continue to concentrate on applying this protocol to both pore and non-pore based detection system. Initial attempts to quantify the signal have proved to be limited by the low resolution of the CCD camera available to the project but are indicative of the size distribution that might be expected from an image generated by the Gaussian envelope of the illumination beam. We would expect, over the final part of the project to be able to show that rapid, solution phase sizing of DNA particles is achievable using our approach. Of particular importance will be the sizing of very large DNA fragments labelled with TOTO or YOYO (we are currently working on

S. pombe chromosomal DNA, 3.5, 4.6 and 5.7 megabase). This is expected to be a significant improvement on current techniques which are unable to either separate (e.g. gel) such large molecules or which place too much shear on the delicate molecules (e.g. flow cytometry) for them to be analysed intact.

B. The Biotechnological Institute (Denmark) has been working on the following tasks:

1. Behaviour of large DNA fragments in free Solution or in gel

We have made a literature survey on electrophoresis of both small and large double stranded DNA molecules both in solution and in gel. All literature records showed that the movement of especially large fragments is very complex. Currently the most popular theory of DNA-movement in gels is the reptating DNA model, describing the movement as snake-like. Importantly, it has been shown that the DNA-molecules move head on, i.e. one of the ends is always preferentially forward. New data have shown that the DNA-molecule in gels not only reptates but also goes through cycles of contraction, fibre collision and elongation. So in summary, the movement of DNA in gels is very complex. With respect to this electrophoresis project, the duration of the cycling in DNA conformations and the elasticity of the large molecules are very important. A basic requirement for a correlation between time, fluorescence signals and physical length is that the molecule is stretched out as it passes the membrane pore.

2. Investigations concerning some of the electrophoretic conditions for the electrophoresis of large DNA molecules through Nucleopore membranes.

We have carried out experiments with electrophoresis of DNA molecules through 6 um thick Nucleopore membranes with pore diameters ranging from 15 to 400 nm and pore densities of approximately 1 to 6 x 10^8 cm^{-2}.
An electrophoresis unit with an electrophoresis chamber (3 x 0.1 cm i.d.) was constructed.
Electrophoresis of bromophenolblue (BPB) was conducted initially in order to study the electrophoretic conditions. Electrophoresis of BPB versus time was linear for at least 5 min. Electrophoresis of BPB was linear at least up to 60 V. Current drift was investigated at constant voltage for at least 15 min. As expected the current was lowered from approximately 100 to 95 uA. The resistances of the membranes were investigated. The same current drift was observed for an electrophoresis tube with or without a Nucleopore membrane. This result was somewhat unexpected, and is not readily explained, but showed that the major contributor to resistance is not over the 6000 nm thick membrane.
DNAs of different molecular weights, 3.6 or 48.5 kD were readily electrophoresed through the 50 nm Nucleopore membrane. This was shown by the polymerase chain reaction. At present it is not known whether there is linearity between voltage and the amount of DNA that passes the Nucleopore membrane. An important result that arose from the DNA experiments was that the DNA presumably is physically intact after passage of the membrane. This was shown by Southern blotting of lambda DNA before and after electrophoresis through the membrane. In summary it is possible to electrophorese a large DNA fragment through the membrane without destroying it physically.

3. Fluorophore labelling of DNA fragments.

The last part of our work has focused and focuses on how to label a large double Stranded DNA fragment. Initially it was thought that it would be possible to anneal fluorophore-labelled oligos to single stranded DNA with concomitant gap-

filling by DNA polymerases in order to generate double-stranded fragments. However, this approach was abandoned as the processivity of DNA polymerases and the possibility of filling out all gaps between oligos are to low.
We have focused on the use of peptide nucleic acids to label double stranded DNA. Peptide nucleic acids have a peptide backbone instead of the sugar phosphate backbone and show better binding properties to DNA than DNA itself. Furthermore the PNA molecule is electrically neutral which is an advantage. We have made initial binding experiments with a fluorescein labelled T8 PNA probe which hybridises to $(dA)_8$ sequences in DNA. That the T8 PNA hybridises to dA-sequences in DNA was confirmed by gel-shift studies with an dA_8 DNA oligo. It has not yet been possible to elucidate good conditions under which the PNA will bind double stranded DNA properly. Further binding experiments will be performed with PNA oligos containing all four nucleotides (ACGT).

C. Optimum Ltd. (Athens) Report

During the first period of work at Optimum, different data analysis scenarios were introduced and theoretically assessed with respect to the range of intensities and speeds of signals expected from the aperture device in order to best predict the hardware requirements of the proposed system.

Optimum has modified the software used by a CCD system with a PCX high resolution video line digitiser employed in the CAMR system. Specifically, the file handling routines were improved allowing data recall and retrospective manipulation of data collected previously in longer experimental ruins. The speed with which the data handling routines could be operated were increased permitting faster analyses in the CAMR studies. A Major advance in this part of the work was achieved through improved data storage compression on disk. In the second phase of the programme, Optimum designed and implemented a hardware independent data analysis package by using the filter driver technique. It was obvious that the data analysis task should run under a graphical computer environment. The popularity and ease of use of MS-Windows made us decide to choose MS-Windows as the operating platform to develop and run the programme.

Two basic principles were employed during the development of the data analysis programme: the event driven programming principle and the Non-modal Programming principle. In order to satisfy the data analysis requirements Optimum selected object orientated languages for software development. This is because object oriented languages offer the ability for event driven and non-modal programme design, which both offer increased user-friendliness and functionality. Various user-friendly tools (virtual instruments) were used to operate the system (scroll bar, option buttons, check buttons, etc...).

The data analysis programme consisted mainly of these parts:

The Filter Driver Module, which performs the raw data storing in packed form according to various data acquisition hardware instrumentation.

The Pre-Processor Module which normalises the data, filters the noise and prepares the data for subsequent data analysis.

The Visual Data Analyser Module (VDU) which performs the data analysis and evaluation.

The Filter Driver handles the data coming from the data acquisition hardware. This filter Driver was necessary because of the vast differences between the different data acquisition instrumentation. The user can easily select the data type

used by the data acquisition system and there is a provision for the use of new data descriptions, which are going to be added in the future months towards the end of the programme., and as newer data acquisition instruments are introduced. The output of the filter driver routines is a standard form of raw data to be stored fro analysis in the future.

The Pre-processor module accepts data in the standard form that the Filter Driver has captured. The programme operator, by visualising these data, can define a set of parameters need to transform the data into a more useful and compact form. The raw data contain a lot of noise and trivial information, which hinder the job of the operator, as they tend to obscure the non-trivial information and slow down the data analysis process accordingly. Furthermore, this trivial information takes up valuable disk and memory space making the analysis of long time observations impossible.
The output of the Pre-processor takes up much less space and allows much faster data analysis without loosing any of the important information taken from the raw data. To ensure data integrity, a set of additional routines has been added.

From the **VDU control panel** the user can select and perform various data analysis tasks by using the pre-processor output for specific observation. The main data analysis tasks are peak height analysis, peak width analysis, peak distribution analysis and time-between-peaks analysis. The basic user interface tools which are used in the VDU are the Chart Window and Spreadsheet Box. In the Chart box the user can display and analyse the data with many enhanced features. In the Spreadsheet box the user can view the actual values of data being analysed. The information is concurrently displayed on both the Spreadsheet box and Chart Window. Each of these two tools can be selected for maximising so that more complicated data can be read more easily. The system offers a set of statistical, mathematical and evaluation functions which may be applied to the data being analysed. Furthermore the user may select a region to zoom in on and analyse. This selection can be made first by selecting two points and then pressing the zoom button. The points can be selected by clicking in the Chart or the Spreadsheet or by clicking once in each. The spreadsheet can be divided into two or four sections, where each section displays a different part of the data, so that the user can monitor the data which is most useful. Moreover, for easier navigation through the data values the Spreadsheet automatically scrolls to the region of data the user selects from the Chart.
The user can set up and control the programme through a series of visual objects, such as radio buttons, buttons, pop-down list boxes, etc.. The user can navigate the whole application by using pointing devices such as a mouse, trackball digitiser, etc.. The keyboard is not necessary to operate the programme eliminating system 'hangs' and making the process more user friendly.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The results of the project so far have shown that individual DNA molecules as small as 50Kb can be followed individually in real-time in solution. Over the remainder of the project the further development of the technique to the analysis of very large (0.1-10Mb) fragments is expected to be of particular commercial interest. The project has been of interest to and assisted by (on a non-participant basis) a major European manufacturer of DNA sequencing equipment.

MAJOR COOPERATIVE LINKS

Full meetings of all participants have been held in Denmark, Athens and the UK over the lifetime of the project and interested commercial organisations in this field have also visited Denmark and the UK on several occasions to discuss possible exploitation and to give technical assistance. In addition, several technical meetings have been held between participants when the need arose or when an opportunity presented itself.

The participants were also collectively invited to present their work at
(i) the EC meeting on Sequencing the Yeast Genomes II and XI (Brugge, Belgium - Oct 91) and at
(ii) the EC meeting on Genome Analysis in the EC (Elounda, Crete - May 91).

PUBLICATIONS

Joint publications

Carr, R., Holmstrom, K., Buchardt, O. and Georgiou, C. (1991) 'Alternative Methods to DNA Sequencing', Abs. EC BRIDGE Meeting on Sequencing the Yeast Chromosomes II and XI, Brugge Sept., 1991, p 90-91.
Perkins, E., Georgiou, C. and Carr, R (1994). Free solution sizing of DNA Fragments, (in prep).

Individual publication

Carr, R. (1991) 'A possible Technique for Solution phase Sequencing of DNA', Abs. EC Int. Meet. on Genome Analysis in the EC, Crete, May 1991.

AREA C:

CELLULAR BIOLOGY

PHYSIOLOGY AND MOLECULAR GENETICS OF INDUSTRIAL MICROORGANISMS

Integration of primary metabolism, secondary metabolism and differentiation in *Streptomyces coelicolor*: A biochemical, physiological and genetical approach (BIOT CT-910255)

COORDINATOR:
A) D.A. HOPWOOD, John Innes Centre, Norwich, GB
PARTICIPANTS:
B) L. DIJKHUIZEN, University of Groningen, NL
C) P. POSTMA, University of Amsterdam, NL
D) C. HARDISSON and J. SALAS, University of Oviedo, E
E) C.J. THOMPSON, Institut Pasteur, Paris, F
F) F. MALPARTIDA, C.S.I.C., Madrid, E
G) D.A. HODGSON, University of Warwick, GB

BACKGROUND INFORMATION

Despite the huge biotechnological importance of streptomycetes as antibiotic producers, and much progress in analysing and manipulating antibiotic biosynthetic genes, molecular knowledge of the regulation of antibiotic production is meagre. Such knowledge should come from an integrated study of a single genetically amenable strain, *Streptomyces coelicolor* A3(2), by a multidisciplinary approach feasible only in this model strain. The project involved six inter-related topics.

(1) To identify factors controlling carbon flux during the switch to secondary metabolism, and to indicate the availability of intermediates of central metabolism for secondary metabolite formation, the routes of glucose catabolism would be established at various growth stages. Cloning genes for regulated enzymes and genetic manipulation of their activity should allow the measurement of relevant fluxes.

(2) Carbon catabolite repression can interfere with antibiotic production. Many metabolic pathways are catabolite repressed, but only in a few enteric bacteria are there insights into the mechanisms. Earlier work had failed to reveal a glucose: PEP-dependent phosphotransferase system in *Streptomyces* and the lack of an apparent role for cAMP in glucose repression indicated that the mechanism could be interestingly different from that in enterics. This was supported by the discovery that a functional glucose kinase gene *(glkA)* was required for glucose repression in *S. coelicolor.*

(3) Streptomycetes were known to accumulate glycogen in a temporally and spatially controlled manner such that it might compete with antibiotics for precursors and energy, or that glycogen could itself be a precursor of other secondary metabolites. It was therefore important to investigate glycogen metabolism in *S. coelicolor.*

(4) The gene cluster encoding the whole actinorhodin biosynthetic pathway had been sequenced. Expression of the biosynthetic genes seemed to be controlled by at least two kinds of elements: a specific transcriptional activator (*act*II-ORF4) within the cluster, and unlinked, pleiotropic, genes that might regulate the biosynthetic genes by targeting more general processes within secondary metabolism. In *S. lividans*, a close relative of *S. coelicolor*, actinorhodin production is virtually unexpressed under non-stress conditions, providing a visual

screen for homologous and heterologous DNA that might trigger production of secondary metabolism in general in streptomycetes.

(5) Earlier work with *S. antibioticus* had suggested changes in specific ribosomal proteins associated with development; the idea was to see whether such changes, if present in *S. coelicolor*, could play a role in metabolic switching.

(6) Understanding of the *Streptomyces* developmental phases leading to antibiotic biosynthesis and morphological differentiation was based largely on genetic studies. We decided to test a complementary approach by using 2D gels to define characteristic patterns of gene expression during development.

OBJECTIVES AND PRIMARY APPROACHES

By diverse approaches, to follow glucose in its journey into the *Streptomyces* cell, through primary metabolism (examining its regulatory consequences) and into storage compounds and secondary metabolites, and to relate this to major developmental and metabolic switches.

RESULTS AND DISCUSSION

1. Glucose metabolic pathways (Groningen)

Measurements of enzyme activities at various growth stages, including the transition from primary to secondary metabolism, indicate that the major routes for glucose metabolism in *S. coelicolor* A3(2) strains M145 and MT1109 are the glycolytic and pentose phosphate pathways; key enzymes of the Entner-Doudoroff pathway were not detected. These data were supported by glucose radiorespirometry in Manchester (A. I. C. Obanye, S. Oliver). Detailed enzyme studies revealed that *S. coelicolor* has a normal set of glycolytic enzymes, several of them sensitive to metabolic regulation, both in exponential growth and during secondary metabolism. The following enzymes are regulated allosterically at the activity level and have been purified to homogeneity and characterised: phosphofructokinase (inhibited by phosphoenolpyruvate and citrate, stimulated by ADP), pyruvate kinase (activated by AMP) and citrate synthase (inhibited by NADH and 2-oxoglutarate). This shows that primary metabolism is regulated, contradicting earlier speculations about deregulated carbon fluxes as essential factors for initiation of secondary metabolism in actinomycetes. These control mechanisms should help careful tuning of central metabolic pathways during balanced growth. The N-terminal amino acid sequence of phosphofructokinase has been determined, allowing cloning of the structural gene.

Growth kinetics, nutrient consumption and production of red/blue pigment (actinorhodin and its congeners) by MT1109 and M145 were studied in batch, fed-batch and continuous cultures. Pigments were produced on depletion of phosphate, ammonium, or trace elements. Ammonium limitation was always accompanied by secretion of pyruvate, indicating that under these conditions the glycolytic pathway suffered a bottleneck at the level of pyruvate. Abundant production of actinorhodin was strictly dependent on medium pH. At pH 4.5-5.5, actinorhodin accumulated intracellularly; at pH 6.0-7.5, actinorhodin was a minor product and instead an extracellular blue pigment, identified as γ-actinorhodin, was the major product.

2. Glucose transport and carbon catabolite repression

(a) Glucose kinase and catabolite repression (Amsterdam, Norwich)

The *glkA* product resembles a family of repressor proteins, but lacks the characteristic helix-turn-helix motif for DNA-binding; the similarity may therefore reflect a common ancestry as sugar-binding proteins. The flux through glucose kinase might have been the critical factor in mediating glucose repression. However, *glkA* deletion mutants expressing either an unrelated glucose kinase from *S. coelicolor* or the *glk* gene of *Zymomonas mobilis* failed to restore glucose repression of the agarase gene *(dagA)*; moreover, over-expression of *glkA* relieved glucose repression of dagA, consistent with a more specific regulatory role for GlkA. These observations are reminiscent of *Saccharomyces cerevisiae*, where hexokinase PII is essential for glucose repression; its role cannot be filled by glucokinase. Agarase production is repressed by several other carbon sources; remarkably, although most of them are not metabolised via glucose kinase, repression of agarase production was relieved in a *glkA* mutant. Similar results were obtained for repression of glycerol kinase. We believe that GlkA plays a central role in carbon catabolite repression in *S. coelicolor* by sensing a signal that reflects either carbon source or energy level, which is then probably transmitted from GlkA to individual genes and operons through interactions with other, as yet unidentified, proteins.

GlkA is ATP-dependent; comparisons with other ATPases led to the identification of an ATP fold in GlkA consisting of five sequence elements. Site specific mutations of the fold were made to affect ATP binding and potentially catabolite repression. The effects of the mutations on enzyme activity and carbon catabolite repression will be assessed. Using a new screening procedure, mutants lacking glucose kinase activity, but still showing glucose repression, were isolated. Future analysis of the various mutants should clarify the role of glucose kinase.

In *S. cerevisiae*, interaction between the glucose transporter(s) and hexokinase PII might play a role in signalling glucose repression. Consequently, a PCR fragment of 850 bp was derived from *S. coelicolor* DNA by using degenerate oligonucleotides corresponding to conserved motifs of eukaryotic hexose transporters; its predicted translation product resembles a putative yeast metabolite transporter. The PCR fragment will be used to inactivate the *S. coelicolor* gene and the effect on sugar transport will be assessed.

Collaborative studies with Colin Smith (UMIST, Manchester) on carbon catabolite repression of the *S. coelicolor gyl* operon showed that *gyl* transcript levels, glycerol kinase activities and the amount of glycerol kinase (measured with an antibody against the *Enterococcus faecalis* enzyme) were regulated entirely at the level of the inducer, glycerol 3-phosphate. Carbon catabolite repression of *gyl* can be explained by modulation of glycerol kinase activity and/or interactions between the repressor of the *gyl* operon, GylR, and glycerol 3-phosphate.

(b) Cyclic AMP (Warwick, Paris)

cAMP is synthesized and excreted by many streptomycetes, yet adenyl cyclase had not been characterized, and its gene *(cya)* had not been identified. Unlike the situation in *E. coli*, cAMP does not influence glucose repression in *S. coelicolor*. Nor could our studies correlate changes in intracellular or extracellular cAMP levels with obvious physiological or morphological processes. However, cAMP levels peaked early in cultures initiated from spores, suggesting a connection with

germination. To clarify the role of cAMP, we characterized the *S. coelicolor* adenyl cyclase enzyme and the corresponding gene.

Adenyl cyclase can be detected in cell-free extracts of *S. coelicolor*. This is different from the situation in enteric bacteria, where loss of cell integrity leads to loss of the vast majority of adenyl cyclase activity. The enzyme prefers Mn^{2+} to Mg^{2+}, but in the presence of the latter it is stimulated by pyruvate. The majority of it is present in the cytoplasmic fraction, but removal of the membrane leads to loss of 50% of total activity.

The *S. coelicolor cya* gene was cloned by complementing an *E. coli cya* mutant. The predicted amino acid sequence identified the enzyme as being most closely related to that of *Brevibacterium liquefaciens*. A mutagenized *cya* gene was constructed and used to disrupt the *S. coelicolor* chromosomal gene. The resulting mutant is partially defective in both spore germination and normal mycelial growth. In liquid cultures or on agar plates, the mutant strain grows as compact mycelial mases with much reduced peripheral filamentation. Addition of cAMP to *cya*$^-$ cultures induces synthesis of at least five proteins and allows normal germination and growth.

3. Storage compound synthesis and metabolism (Oviedo, Norwich)

(a) Biochemical and physiological studies.
Most of this work used *S. antibioticus*, which is particularly active in glycogen deposition. Shift-down conditions (N-limited, glucose excess) that activate glycogen synthesis have been defined. Very young hyphae cannot be induced in this way, indicating developmental requirements that are currently being studied. Induction is at the level of transcription. The major enzymes of glycogen synthesis co-fractionate with glycogen particles, apparently together with a novel glycogen protein-priming activity that may provide the oligoglucan primer for extension by glycogen synthase. Levels of glycogen and trehalose after subjecting down-shifted cells to osmotic shock, temperature changes and nutritional upshift implicate trehalose as a generalised protecting agent, and glycogen as a significant carbon source for trehalose synthesis only during osmotic adjustment.

(b) Morphological studies.
Electron microscopy of colonies of *S. coelicolor* wild-type and 13 representative morphological mutants suggested that the two phases of glycogen deposition that occur during normal development are differently controlled. Phase I, at the interface of substrate and aerial mycelium, is virtually absent from *bldA, bldB, bldC, bldG* and *bldH* mutants (but abundant in *bldF*), even when their normally aerial mycelium-defective phenotype is corrected by growth on a different carbon source. In this respect, phase I glycogen behaves rather like secondary metabolites, which are also unconditionally absent from most of these mutants. On the other hand, phase II glycogen deposition (in immature spores) is independent of *bld* genes (where this is testable), but instead shows some dependence, direct or indirect, on the *whiG* gene, which encodes a putative sigma factor that initiates sporulation.

(c) Genetic studies.
PCR-mediated amplification, using primers based on conserved regions in relevant enzymes of glycogen metabolism, has led to the cloning of two branching enzyme genes (*glgB*I and *glgB*II), one (perhaps two) candidate ADP glucose pyrophosphorylase genes *(glgC)*, and two candidate glycogen phosphorylase genes *(glgP)*. At least three (perhaps four) regions of the genome contain *glg* genes. This situation is more reminiscent of the situation in higher eukaryotes than that in other bacteria, which generally have a single set of clustered *glg* genes. At least for

the *glgB* genes, the duplication correlates with the two phases of glycogen deposition: disruption of *glgB*I abolished glycogen production (resulting in unbranched glucan deposits) at phase I and had no effect on phase II, while disruption of *glgB*II had the reciprocal effect, leaving phase I deposition undisturbed but causing lack of branching in phase II.

4. Regulatory cascades that lead to antibiotic synthesis (Madrid, Norwich)

Stationary phase production of actinorhodin and undecylprodigiosin appears to be mediated through transcriptional activation of their respective pathway-specific activator genes, *act*II-ORF4 and *redD*, which normally occurs during transition phase; their artificial expression in exponential growth causes early and overproduction of the respective antibiotic. *act*II-ORF4 transcription also increases after nutritional shiftdown, correlating with the appearance of ppGpp; in contrast, transcription of *rrnD* (a rRNA gene set), *glkA* and *afsR* (a pleiotropic regulator of actinorhodin and undecylprodigiosin production) decrease markedly, both in transition phase and after nutritional shiftdown. Lack of transcription of *redD* after shiftdown suggests that, if ppGpp plays a role in determining the onset of antibiotic production, it is not always sufficient. A *Streptomyces* ppGpp synthetase gene was cloned with a probe based on the N-terminal sequence of a synthetase from *S. antibioticus*. The chromosomal copy of the *S. coelicolor* gene is being disrupted and the effects on antibiotic production assessed.

The single TTA codon in *act*II-ORF4 is responsible for dependence of actinorhodin production on *bldA*, which encodes the $tRNA^{TTA}_{Leu}$. In these studies we detected no conditions in which *bldA* was particularly poorly transcribed or in which the translation of TTA codons was limiting for gene expression. The role of *bldA* in determining the onset of actinorhodin production therefore remains problematical. Transcription of *act*II-ORF4 is *bldA*-independent, but *redD*, which is TTA-free, is not transcribed in *bldA* mutants; presumably an additional TTA-containing gene (or genes) is needed for *redD* transcription, and such genes are being sought.

Expression of *act*II-ORF4 in *E. coli* yielded inclusion bodies that were solubilized to yield a >90% pure preparation. This was used in gel-shift assays with two *act* DNA fragments that induce an Act^- phenotype at high copy number: these are the intergenic regions *act*VI-ORF1/*act*VI-ORFA and *act*I-ORF1/*act*III. The *E. coli*-derived protein, and extracts of *S. coelicolor* and *S. lividans* that contained ActII-ORF4, all failed to cause band shifts. Thus ActII-ORF4 probably lacks DNA binding properties on its own; other elements may be needed for specific interaction with the DNA.

In similar experiments, RedD was over-expressed in *E. coli* and *S. coelicolor* as a His-tagged fusion protein After purification from *E. coli*, the protein was used to raise antibodies. Western analyses failed to detect RedD in *S. coelicolor* during undecylprodigiosin production. Filter retention and gel-shift assays using RedD from both sources failed to reveal any DNA-binding activity *in vitro*, so RedD, like *Act*II-ORF4, may act in concert with at least one additional protein, which is currently being sought.

RNA polymerase holoenzymes from transition phase *S. coelicolor* that recognise the *act*II-ORF4 and *redD* promoters differ chromatographically from the major holoenzyme. Whether this reflects the acquisition of an alternative σ factor, covalent modification of the holoenzyme, or the degradation or processing of the

major σ factor (σ^{hrdB}) to a form still recognising the *redD* and *act*II-ORF4 promoters is being analyzed by reconstitution experiments.

The *abaA* locus (antibiotic biosynthesis activator) from *S. coelicolor*, originally cloned by its ability to induce blue pigment production in *S. lividans*, contains three ORFs: ORFA, B (complete) and C (truncated). Insertional inactivation of ORFA or B (but not ORFC) in *S. coelicolor* generated mutants unable to produce actinorhodin, undecylprodigiosin or CDA, but that still produced methylenomycin. Cloning ORFA at high copy number in *S. coelicolor* abolished actinorhodin, prodigiosin, CDA and spore production, a phenotype resembling that of *bldA* mutants. The DNA sequence of the intergenic region between ORFB and C was responsible for inducing actinorhodin production in *S. lividans*.

To analyse other pleiotropic regulatory elements, two blue colonies were isolated from a genomic library of *S. antibioticus* DNA cloned in *S. lividans*. One contained the 3' end of an ORF (ORF1) and a complete second ORF (ORF2). The deduced product of ORF1 resembles that of *metB* of *E. coli*, while ORF2, which is responsible for stimulation of actinorhodin production, resembles transcriptional regulators of the LysR family. Interestingly, ORF2 contains a TTA codon and so depends on *bldA* for expression. The gene was named *abaB*. It is present in many *Streptomyces* spp. including *S. lividans* and *S. coelicolor*.

A pair of divergently transcribed genes (ORF10/ORF11) was identified immediately to the right of the known *act* cluster of *S. coelicolor*. ORF11 apparently encodes a ketoreductase (similar to the *act*III product) and the ORF10 product resembles transcriptional regulators of the LysR family. ORF11 can complement the *act*III mutation only when ORF10 is also present at high copy number. Both genes are well conserved in *S. lividans*, and insertional inactivation or chromosomal deletion of ORF10 in *S. lividans* caused actinorhodin overproduction. The ORF10 protein was overexpressed and purified from *E. coli*; using band-shift and footprinting assays, a target was found within or close to the transcription start site for ORF10, but attempts to find other putative targets in the *act* cluster were unsuccessful.

A genomic library from *Streptomyces fradiae*, screened in *S. lividans* for actinorhodin production, yielded a DNA region encoding two divergent ORFs separated by a stretch of DNA that produces a 132 nt long transcript. ORF1 encodes a 99 amino acid polypeptide for which no homologous proteins were found in the database. The truncated ORF2 encodes the carboxyl end of an apparent S-adenosyl-L-homocysteine hydrolase. Subcloning showed that the activating properties of the cloned DNA were due to the intergenic region transcribed from its own promoter, and also from a promoter upstream of ORF1. Both transcripts terminated at a predicted thermodynamically stable secondary structure. Synthesis of the 132 long transcript coincided temporally with that of *act*III, whereas that of the ORF1 transcript did not, consistent with the subcloning results and pointing to a direct role for the 132 nt transcript in activating actinorhodin production. The region of the ORF1 mRNA that includes the ribosome-binding site and translation initiation codon is complementary to the 132 nt long transcript. This might therefore act as an antisense RNA. A corresponding clone was isolated from *S. lividans*. Again, an ORF1 was separated from a truncated ORF2 by a stretch of DNA encoding a shorter, 106 nt long, RNA. Complementarity between the same region of the ORF1 mRNA and the 106 nt long transcript was found. Transcription initiation and termination patterns were very similar to those for the *S. fradiae* clone, and again transcription from the putative antisense RNA promoter coincided with that of *act*III, whereas transcription from the ORF1

promoter started much earlier. We propose the name *micX* for the genes encoding these putative antisense RNAs.

5. Changes in ribosome structure during development (Oviedo)

To monitor the kinetics of ribosomal protein synthesis during growth, *S. coelicolor* cultures were pulse-labelled with ^{35}S-methionine at early exponential phase, transition from primary to secondary metabolism, and stationary phase in minimal medium. Total cellular proteins were extracted, separated by 2D gel electrophoresis and revealed by autoradiography. Most of the ribosomal proteins were resolved as single spots. The specific rate of synthesis of most of them was higher at early stages of growth, decreasing with the transition phase. Interestingly, two protein spots showed a different pattern: their synthesis drastically diminished after the transition phase and was undetectable later. One of them corresponded in its molecular weight and isoelectric point to L7/L12 (L7 is an acetylated form of L12) from *E. coli*. The gene for L7/L12 is normally cotranscribed with the L10 gene. A previously isolated DNA fragment of *S. antibioticus* encoding these two proteins was used to clone the *S. coelicolor* homologues. Sequencing revealed three complete ORFs and the 3' end of an incomplete ORF. The deduced amino acid sequences of two of the ORFs clearly resembled L7/L12 and L10 from different organisms. The other two ORFs showed no similarity with known proteins. The intergenic region between one of these ORFs and the L10 gene contains a sequence that can be folded into a stem-loop structure resembling a region of *S. ambofaciens* 23S ribosomal RNA and is thus likely to be a L10 binding site for translational regulation. To generate polyclonal antibodies against L7/L12 and L10, the 3' end of each gene was fused to *lacZ*, over-expressed in *E. coli* and the fusion protein purified by affinity chromatography. The antibodies recognised proteins in positions corresponding to the two proteins that were differentially expressed during growth. Promoter-probing with pIJ4083 (containing the promoterless *xylE* gene) indicated four potential promoters of different strength in the sequenced region. The 5' ends of the corresponding transcripts were mapped by primer extension. At least one of the promoters showed well conserved *E. coli* 'consensus'-like -35 and -10 regions.

The L10 and L7/L12 genes were expressed in *S. lividans* after creation of an *Nde*I site at each start codon for cloning into the expression vector pIJ6021. Controlled expression of both proteins from the *tipA* promoter will enable us to study translational regulation of this operon. Work in progress is focused on constructing *xylE* translational fusions to the 3' ends of the L10 and L7/L12 genes. These will be stably integrated into the *Streptomyces* chromosome using integrative vectors containing the øC31 *att* site to test whether overproduction of L10 or L7/L12 can switch off the synthesis of both ribosomal proteins.

6. Variations in proteins regulated by global metabolic switches (Paris, Oviedo)

Rates of synthesis of the majority of *S. coelicolor* proteins were measured in different phases of growth, physiological conditions, or mutant backgrounds. Liquid cultures were pulse-labelled with ^{35}S-methionine and the labelled proteins separated on 2D gels. Computer analysis was used to quantify protein spots on autoradiograms, and to define clusters of proteins with similar patterns of expression. We have constructed a database describing the rates of synthesis of more than 500 *S. coelicolor* protein spots as a function of the growth curve and identified gene products associated with actinorhodin biosynthesis, the heat shock stress response and the ribosome.

Under our conditions, at least four different phases can be identified in the growth curve: 'exponential growth', 'hesitation', 'secondary growth' and 'stationary'. A transitory hesitation in growth occurs late in the exponential phase despite the absence of significant changes in the levels of carbon (maltose) and nitrogen (glutamate) in the medium. We therefore believe that this 'hesitation' phase may be triggered by **localized** nutritional imbalances within the mycelial pellets. During this period of several hours, rates of protein synthesis decrease by more than 95%. After this, a short secondary phase of growth takes place before stationary phase. The four phases are characterized by remarkably different patterns of gene expression: between one-third and more than half of the spots change by a factor of >10. These unexpectedly complex changes in gene expression during growth may reflect its unusual developmental programme leading to secondary metabolism or morphological differentiation; such extensive changes in 2D gel patterns have not been reported for *E. coli*. We are interested in the 'hesitation' phase because it is at this time that our 2D protein database has revealed three important events: activation of actinorhodin production genes, suppression of ribosomal protein synthesis, and activation of a heat shock stress response.

Presumptive antibiotic biosynthetic gene products (>20) were identified by their absence in a mutant in which the actinorhodin activator (*act*II-ORF4) had been disrupted (provided by the Madrid group); preliminary characterization of the rates of synthesis indicates that, while all gene products first appear during the hesitation phase, they have different kinetics. Synthesis of most of the 27 ribosomal proteins detected continued throughout all phases of growth in liquid cultures, but their relative rates of synthesis decreased sharply with respect to overall protein synthesis during the 'hesitation phase'. A simple, global 'heat shock stress response', as described in other, nondifferentiating eubacteria, was not observed in *S. coelicolor*. Instead, the response to heat, or presumably other stress conditions such as antibiotic biosynthesis, was developmentally regulated. About half of the 23 heat-induced protein spots were also induced by stress conditions during the 'hesitation' phase. The growth phase also determined the kinetics of induction for virtually all HSPs; some HSPs were induced only at a certain stage of 'physiological differentiation'.

MAJOR SCIENTIFIC BREAKTHROUGHS

Identification of glycolysis and the pentose phosphate pathway (not the Entner-Doudoroff pathway) for glucose assimilation, and proof of **regulation** of glucose metabolism; cloning of some of the genes for flux control analysis. Implication of the conversion of actinorhodin to γ-actinorhodin in antibiotic export. Proof of a specific role for the glucose kinase protein in carbon catabolite repression. Cloning of the adenyl cyclase gene and its implication in differentiation. Discovery of a protein-primed glycogen synthase and demonstration of differential control of two distinct phases of glycogen accumulation. Purification of specific activator proteins for two antibiotic pathways and characterisation of several novel pleiotropic activator genes; evidence for an antisense RNA in control of antibiotic pathway genes. Cloning of two ribosomal protein genes and evidence for their differential regulation. Establishment of a database for 500 *S. coelicolor* proteins on 2D gels; proof that developmental phases and stress conditions modulate distinct families of proteins.

MAJOR COOPERATIVE LINKS

All participants, two associated groups (at the University of Glasgow and UMIST, Manchester, GB), and representatives of up to six commercial companies, met together four times, with extensive exchange of information and strains (the latter also throughout the programme). Co-ordinating visits by at least one member of A to all other groups, and many other bilateral visits. Specific collaborations, often with exchange of personnel, took place for: glucose metabolism (B with Glasgow and Manchester), actinorhodin biosynthesis (A,B,F), catabolite repression (A,C), cyclic AMP (E,G), storage compounds (A,D), antibiotic regulation (A,F), ribosome structure (D,E), protein variation under global regulation (D,E), vector development and provision (A,E,F).

PUBLICATIONS

Joint publications

Blanco, G., Rodicio, M. R., Puglia, A. M., Méndez, C., Thompson, C. J and Salas, J. A. (1994) Synthesis of ribosomal proteins during growth of *Streptomyces coelicolor*. Mol. Microbiol. 12:375-385.

Bystrykh, L. V., Dijkhuizen, L., and Obanye, A. I. C. Alterations in primary metabolism upon growth cessation and initiation of secondary metabolism in *Streptomyces coelicolor* A3(2). (In preparation).

Bystrykh, L. V., Dijkhuizen, L., Fernández-Moreno, M. A., Malpartida, F., and Hopwood, D. A. Production of blue pigments by *Streptomyces coelicolor* A3(2). (In preparation).

Danchin, A., Pidoux, J., Krin, E., Thompson, C. and Ullmann, A. (1993) The adenylate cyclase catalytic domain of *Streptomyces coelicolor* is carboxyl-terminal. FEMS Microbiol. Letts. 114:145-152.

Duchêne, A.-M., Kieser, H.M., Hopwood, D.A., Thompson, C.J. and Mazodier, P. (1994) Characterization of two groEL genes in *Streptomyces coelicolor* A3(2). Gene. **144**: 97-101.

Fernández-Moreno, M.A., Martínez, E., Caballero, J.L., Ichinose, K., Hopwood, D.A. and Malpartida, F. (1994) DNA sequence and functions of the *act*VI region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor* A3(2). J. Biol. Chem. **269**, 24854-24863.

Individual publications

Alves, A., Dijkhuizen, L. Purification and characterization of an ATP-dependent phosphofructokinase from *Streptomyces coelicolor* A3(2). (In preparation).

Angell, S. (1992) The glucose kinase gene of *Streptomyces coelicolor* A3(2) and its role in glucose repression. Ph.D. Thesis, University of East Anglia, Norwich, GB.

Angell, S., Schwarz, E. and Bibb, M.J. (1992) The glucose kinase gene of *Streptomyces coelicolor* A3(2): its nucleotide sequence, transcriptional analysis and role in glucose repression. Mol. Microbiol. 6:2833-2844.

Angell, S., Lewis, C.G., Buttner, M.J. and Bibb, M.J. (1994) Glucose repression in *Streptomyces coelicolor* A3(2): a likely regulatory role for glucose kinase. Mol. Gen. Genet. **244**: 135-143.

Blanco, G., Parra, F., Méndez, C. and Salas, J.A. (1992) The nucleotide sequence of the L10 equivalent ribosomal protein gene of *Streptomyces antibioticus*. Nucleic Acid Research, 20:5223.

Bruton, C.J., Plaskitt, K.A. and Chater, K.F. Tissue-specific isoenzymes of glycogen branching enzyme in a mycelial prokaryote (In preparation).

Champness, W. and Chater, K.F. (1994) Regulation and integration of antibiotic production and morphological differentiation in *Streptomyces* spp. in P. Piggott et al. (eds), Regulation of Bacterial Differentiation, pp. 61-92. American Society for Microbiology, Washington, D.C.

Fernández-Moreno, M.A., Martínez, E., Martínez-Costa, O., Martín-Triana, A.J. and Malpartida, F. (1993) Workshop on Molecular Genetics in Modern Biotechnology. Mallorca, Spain, October 1993.

Gramajo, H.C., Takano, E., and Bibb, M.J. (1993) Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. Mol. Microbiol. 7:837-845.

Kwakman, J.H.J.M., Debije, C., and Postma, P.W. (1992) Catabolite repression in *Streptomyces coelicolor* A3(2). Fifth ASM Conference on Genetics and Molecular Biology of Industrial Microorganisms, October 11-15, 1992, Bloomington, USA, p.13.
Kwakman, J.H.J.M. and Postma, P. (1994) Glucose kinase has a regulatory role in carbon catabolite repression in *Streptomyces coelicolor*. J. Bacteriol. 176:2694-2698.
Kwakman, J.H.J.M. and Postma, P.W. (1994) Role of glucokinase in carbon catabolite repression in *Streptomyces coelicolor*. Abstracts 127th Meeting of Society for General Microbiology, 6-7 January 1994. Warwick, U.K., p.33.
Martín-Triana, A.J. (1994) Caracterizacion de dos sistemas geneticos implicados en la regulación de la síntesis de antibióticos en *Streptomyces*. Ph.D. Thesis, Universidad Autónoma de Madrid, Madrid, Spain.
Parra, F., Blanco, G., Alonso, J.M.N., Balbin, M., Méndez, C. and Salas, J.A. (1992) Cloning and sequence of a gene encoding the L7/L12 ribosomal protein equivalent of *Streptomyces antibioticus*. Gene 118:127-129.
Parro, V. and Mellado, R.P. (1994) Glucose effect on agarase overproduction by *Streptomyces*. Gene. **145**: 49-55.
Plaskitt, K.A. and Chater, K.F. (1994) Influences of developmental genes on localized glycogen deposition in colonies of a mycelial prokaryote, *Streptomyces coelicolor* A3(2): a possible interface between metabolism and morphogenesis. Phil. Trans. Roy. Soc. B, in press.
Puglia, A.M., Vohradsky and Thompson, C.J. (1994) Developmental and thermal regulation of the heat shock response in *Streptomyces coelicolor*. Mol. Microbiol. In preparation.
Romero, N.M., Parro, V., Malpartida, F., and Mellado, R.P. (1992) Heterologous activation of the actinorhodin biosynthetic pathway in *Streptomyces lividans*. Nucleic Acids Research 20:2767-2772.
Romero, N.M., Parro, V. and Mellado, R.P. (1994) Expression of a heterologous gene activating actinorhodin biosynthesis in *Streptomyces* lividans and *Streptomyces coelicolor*. FEMS Microbiol. Lett. 116:301-306.
Romero, N.M., and Mellado, R.P. (1994) Homologous activation of the actinorhodin biosynthetic pathway in *Streptomyces lividans*. In preparation.
Scheu, A.K. (1994) Aislamiento y caracterización del gen *abaB* de *Streptomyces antibioticus* ATCC11891 que codifica un activador pleiotrópico de la biosíntesis de antibióticos en *Streptomyces lividans*. Ph.D. Thesis, Universidad Autónoma de Madrid, Madrid, Spain.
Servín-González, L., Jensen, M.R., White, J. and Bibb, M.J. (1994). Transcriptional regulation of the four promoters of the agarase gene (dagA) of *Streptomyces coelicolor* A3(2). Microbiol. **140**: 2555-2565.
Takano, E., Gramajo, H.C. Strauch, E., Andres, N., White, J. and Bibb, M.J. (1992) Transcriptional regulation of the redD transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). Mol. Microbiol. 6:2797-2804.
Takano, E. (1993) ppGpp and antibiotic production in *Streptomyces coelicolor* A3(2). Ph.D. Thesis, University of East Anglia, Norwich, GB.
Takano, E., White, J., Thompson, C.J. and Bibb, M.J. Construction and application of a multicopy thiostrepton-inducible expression vector for *Streptomyces*. (In preparation)
Takano, E. and Bibb, M.J. (1994) The stringent response, ppGpp and antibiotic production in *Streptomyces coelicolor* A3(2). Actinomycetol. **8**: 1-16.
Takano, E. and Bibb, M.J. Growth phase-dependent gene expression in *Streptomyces coelicolor* A3(2). (In preparation)
van Wezel, G. (1994) Transcriptional regulation of translational genes in *Streptomyces coelicolor* A3(2). Ph.D. Thesis, Rijksuniversiteit, Leiden, NL.

The production and recovery of biotechnologically important proteins from the yeast *Saccharomyces cerevisiae* (BIOT CT-900165)

COORDINATOR:
A. C. HADFIELD, Univ Leicester, Leicester, GB
PARTICIPANTS:
B. D. PIOLI, ZENECA Pharmaceuticals, Macclesfield, GB
C. W.E. LANCASHIRE, Whitbread & Co, Luton, GB
D. P.A. MEACOCK, Univ Leicester, Leicester, GB
E. C. NOMBELA, Univ Complutense, Madrid, E
F. F. del REY, Univ Salamanca, Salamanca, E
G. J. DOMMES, Eurogentec SA, Liège, B

BACKGROUND INFORMATION

Yeasts are important organisms for both classical and modern biotechnology. In particular, *S. cerevisiae*, as the most highly characterised yeast (and eukaryote in general), has a major role. As with all host organisms, however, its performance in the production and recovery of biotechnologically important proteins does have certain limitations. The following aspects were identified as ones were further study would have particular benefits.

(i) Growth-phase responsive gene expression

High level production of a foreign protein from yeast imposes a substantial energy demand that can reduce the overall growth rate of a culture. This may result in outgrowth of the culture by subclones that have undergone spontaneous genetic rearrangement leading to elimination or inactivation of the expression construct. The net result is a reduced product yield from the culture. Also, if the product is toxic to the cell, then early expression will result in loss of culture viability. Thus, it may be advantageous to control the time of expression of heterologous genes during fermentation. For these reasons it would be useful to have available gene promoters that show a naturally regulated pattern of expression during the course of a yeast fermentation. In particular, use of promoters that become activated as a yeast batch culture transitions from rapid exponential growth through aerobic, non-fermentative growth and into stationary phase would facilitate induced expression of heterologous proteins after accumulation of biomass. Such promoters would not require media changes or the addition of specific chemicals in order to induce expression; procedures that can be costly to effect on an industrial scale.

(ii) Protein secretion performance

The protein secretion pathway of yeast is fundamentally the same as that of higher eukaryotes, with many of the associated functions highly conserved. Many of the hormones, growth factors, enzymes, etc. with potentially important biotechnological applications are produced via the secretion pathway in their native producer cells. Such products can be produced biotechnologically in yeast, which provides the necessary post-translational processing. However, it has been found that the secretory pathway of *S. cerevisiae* does not secrete many recombinant proteins efficiently or to a high capacity. In addition, secretion is a preferred route for recovery of recombinant proteins from yeast, as not many host proteins are

secreted — which is a help in subsequent purification — and yeast cells, with their rigid cell walls, are difficult to break open — particularly on a large scale. Thus, improvements in secretion performance are highly desirable.

(iii) Cell wall manipulation

The definition of functions controlling cell integrity and cell wall stability during growth should contribute substantial basic knowledge to be applied for improving the performance of yeast cells in biotechnological processes. The release of proteins (both cytoplasmic and secretory) from yeast cells, a bottleneck in yeast biotechnology, is affected by cell wall dynamics among other factors. The characterization of genes and functions controlling the composition of the cell wall and its osmotic stability represents a useful strategy to manipulate genetic elements that can influence the release of protein from yeast cells. Thus, a less dense cell wall may be favourable to the release of large secretory proteins, whilst inducible autolysis would provide a useful means of releasing cytoplasmic proteins from yeast cells.

OBJECTIVES AND PRIMARY APPROACHES

(i) Analysis of growth-phase responsive genes

Previous work (participant D) had led to the isolation of a set of six cDNA clones representing genes induced as batch cultures, growing on the industrial substrate molasses, reached the end of rapid exponential growth and moved towards stationary phase. Four of these cDNAs had been identified by genetic analysis and DNA sequencing namely: *HSP26* encoding the 26kDa heat shock protein; *HSP12* a new gene encoding a small 12kDa heat shock protein; HXK1 encoding the glucose-repressible hexokinase; *THI4* a new gene encoding a thiamine biosynthetic enzyme. Two genes represented by cDNAs #10 and #15 were uncharacterised.

This project has sought to identify the physiological conditions and mechanisms that cause induction of these genes, and to characterise the structures of the genes corresponding to cDNAs #10 and #15. The investigation aimed to examine expression of the genes in various growth regimes and to dissect their molecular structures. Their biotechnological usefulness would be examined, especially with respect to the *HSP12* promoter, in studies of the production of recombinant trout growth hormone, somatotropin.

(ii) Protein secretion performance

A genetic approach was used to isolate yeast host mutants that exhibited enhanced secretion of recombinant proteins. Such mutants identify and define bottlenecks in the secretion pathway. Heterologous secretory reporters were chosen that were either inefficiently secreted or demanded high capacity secretion by the host cell. Chemical mutagenesis was used to generate mutants that could be identified initially on plate screens and then subsequently by more accurrate quantitative assays.

In addition, production of recombinant trout somatotropin via secretion from yeast was evaluated; also, the secretory β-galactosidase of *Aspergillus niger* was expressed by brewing yeast and examined for fermentation of lactose in dairy waste to ethanol.

(iii) Functions governing cell wall structure and integrity

Genes coding for enzymes able to hydrolyze yeast cell wall compounds will be cloned and characterized. The processing and secretion of these enzymes will be studied, in order to facilitate the production of heterologous proteins. The expression of the genes will be manipulated in order to achieve changes in cell wall dynamics that might enhance the release of periplasmically located proteins.

Genes governing cell wall integrity, which are implicated in a temperature-sensitive autolytic phenotype, will be cloned and characterized. Controlled expression of such genes will be monitored for effects on product secretion/recovery. Autolytic mutant/cell-wall-modified strains will be evaluated for recovery of intracellularly proteins.

RESULTS AND DISCUSSION

A. Analysis of growth-phase responsive genes

1. Studies of Gene structure, location and function

CHEF gel analysis showed the single-copy genes to be located on various chromosomes; *HSP12*, Chr. VI; *HSP26*, Chr.II; *THI4*, Chr. VII.

Nucleotide sequencing showed that the predicted product of cDNA #15 had homology with the family of hexose transport proteins, notably those encoded by *S. cerevisiae* genes *HXT1, HXT2, GAL2.* These are twelve-membrane spanning domain sugar transporters of approx. 65kDa molecular mass. Subsequently it was shown that cDNA #15, also called ESP65, is the recently described *LGT1/RAG1/ (HXT3)* gene encoding a low affinity glucose transporter.

Nucleotide sequencing of the 596bp cDNA #10, also called *ESP30/MOL2*, showed it to contain an incomplete ORF with sequence homology to the *nmt1* gene, of the fission yeast *Schizosaccharomyces pombe*, that is regulated by thiamine and encodes an enzyme of the thiamine biosynthetic pathway. Subsequently, cDNA #10 was shown to derive from a thiamine-regulated gene, see below, that has now been designated *THI5*.

Cloning and analysis of the cDNA #10 gene was complicated by the discovery of at least three homologous sequences within the yeast genome, as evidenced by multiple bands which hybridised to the cDNA on Southern blot analysis of digested genomic DNA. Similarly, hybridisation of cDNA #10 to separated chromosomes (CHEF gel) gave positive signals to chromosomes II, III, VI and XIV.

Genomic clones that hybridised to cDNA #10 were recovered from a λEMBL3-yeast DNA library, and analysis carried out on clones representing two different genomic loci. Hybridisation and sequencing revealed that each carried a region homologous to the incomplete ORF on the cDNA clone, and to the *Sz. pombe nmt1* gene. Use of fragments of unique sequence DNA from these genomic clones as hybridisation probes showed that one derived from Chromosome II and the other from Chromosome VI. The cDNA-homologous region on the Chr.VI clone, was the carboxyl-terminal part of a 1023 bp ORF. The 341aa product of this ORF showed 61% aa sequence identity (78% similarity) to the 347-residue *nmt1* protein. On the Chr.II clone the cDNA-hybridising region, again was homologous to the *Sz. pombe nmt1* gene, with 45% identity (66% similarity).

Comparison of the DNA sequences immediately downstream of these ORFs showed that the Chr VI clone matched the cDNA #10 3'-end exactly for 124 nucleotides, whereas the Chr.II sequence did not; therefore the cDNA #10

derived from the Chr.VI gene. Consistent with this, use of the Chr.VI gene 3'-region as a hybridisation probe detected a thiamine-regulated mRNA, of the same size as that detected using the cDNA (see below); in contrast no mRNA species was detected using the 3'-region of the prematurely terminated ORF on Chr.II. All of these data point to the conclusion that this gene has undergone duplication and evolution to a state where the Chr.VI copy has remained active and the Chr.II copy has become a pseudogene. We have designated the active Chr.VI gene as *THI5*.

A strain with a disruption of the *THI5* gene on Chr.VI. showed a thiamine auxotrophic phenotype, restorable by supplementation with the thiamine precursor hydroxymethyl-pyrimidine. This is consistent with it being the only member of the 'gene family' expressed to give a functional product, namely an enzyme of the biosynthetic pathway of the pyrimidine-moiety precursor.

Curiously, each of these two *nmt1*-like ORFs lies adjacent to another sequence that again occurs in multiple copies in the yeast genome. These sequences are related to the right-most open reading frame of Chromosome III, YCR107W, that is thought to encode an aryl alcohol dehydrogenase. These YCR107W homologues also cross-hybridised to cDNA #10. The YCR107W homologue on Chr.II comprises three short ORFs, all out of frame with each other, but together occupying a region of DNA approximately equal in length to YCR107W. The sequence on Chr.VI contains two short ORFs, again out of frame. It would therefore appear that both of these DNA sequences are pseudogenes related to the YCR107W gene on Chr.III. Use of DNA fragments from these Chr. II and Chr.VI regions failed to detect any homologous mRNA species. Therefore it cannot be concluded whether YCR107W, or either of these homologues, is expressed.

2. Expression studies

Expression of the set of six genes was investigated during batch growth of strain S288c on various substrates, such as glucose (fermentable, catabolite repressing), raffinose (fermentable, non-catabolite repressing), glycerol/ethanol/lactate (non-fermentable, non-catabolite repressing) and on the complex industrial substrate molasses. Two, *LGT1* and *HXK1*, were subject to strong glucose repression, being only expressed when glucose was exhausted or during growth on non-repressing substrates. These observations were extended by examining expression in strains carrying mututations in the glucose/catabolite repression system. Results confirmed *HXK1* and *LGT1* to be subject to carbon catabolite repression via the products of the *TUP1, CYC8* and *SNF1 genes.*

The heat shock genes, *HSP12* and *HSP26*, showed expression during logarithmic growth on aerobically metabolised non-fermentable substrates, and sustained stationary phase induction in all media. Although *HSP26* expression was influenced to some extent by catabolite repression, it was, like *HSP12*, clearly under cAMP control as evidenced by constitututive expression in a *cyr1* mutant and failure to induce in a *bcy1* strain. The promoter regions of both genes contain a DNA sequence which has been shown to mediate transcriptional regulation in response to cAMP. Both genes too, were induced by a variety of physiological stresses including heat shock, alcohol shock, low pH and osmotic shock. All of these treatments affect intracellular pH and ionic balance. It is possible that the genes respond to cAMP-PK via changes in these parameters.

The remaining two genes, *THI4* and *THI5*, were expressed transiently and only in molasses medium. Experiments analysing the effects of supplementing molasses medium showed that production of the *THI4* and *THI5* mRNAs was completely

repressed by exogenous thiamine; *THI4* encodes an enzyme of the biosynthetic pathway leading to the other thiamine precursor hydroxyethyl-thiazole. A twofold rise in intracellular thiamine levels, brought about by rapid and efficient uptake of exogenous vitamin, was sufficient for complete repression of these genes. The time of induction of these genes could be controlled by manipulation of the amount of thiamine supplied in the growth medium; these promoters have considerable attraction as candidates for regulatable expression systems.

These studies have revealed that stationary phase induction occurs through a variety of regulatory mechanisms. Genes and promoters have been characterised that respond to

(i) glucose catabolite repression, *HXK1* and *LGT1*,
(ii) stress and starvation through cAMP, *HSP12* and *HSP26*,
(iii) thiamine depletion, *THI4* and *THI5*.

3. Performance of fermentation-stage-specific expression directed by the HSP12 promoter for the production of trout somatotropin in various fermenter conditions

An expression cassette was constructed in which the cDNA sequence of the trout somatotropin (without its signal sequence) was placed under the control of the *HSP12* promoter. Different yeast strains were transformed by this cassette cloned in a 2μm-based high copy yeast vector. Initial screening in shake flask experiments allowed the production of 4 mg/l of somatotropin.

Several strains were evaluated in fermenter conditions and BJ2168, a protease deficient strain, proved to be the best producer. Expression of the recombinant somatotropin was observed after glucose exhaustion. A transient expression could be induced by heat shock during the exponential growth phase. Heat shocks did not improve expression in stationnary phase. Comparison of different media allowed an increase in biomass accumulation, but the yield on a per cell basis was not improved. Fed batch fermentations were also performed to assess the possibility to induce expression by other metabolic stress, but lower expression levels were obtained. Optimised batch fermentation conditions allowed the production of recombinant trout somatotropin at a level of 100 mg/l. However, the protein was highly insoluble, found in the cells as inclusion bodies.

B. Protein secretion performance

1. S. cerevisiae *mutants exhibiting enhanced secretion of mini-proinsulin*

Three mutants have been isolated that show increases in yield of highly-expressed, secreted mini-proinsulin over the wild-type control of 3-fold for #1, 2-fold for #4, and 2-fold for #9.

Northern blots showed that these mutants appeared to have similar abundances of mini-proinsulin mRNA in their cells — indicating that the gains are in secretion pathway efficiency rather than in expression level.

An alternative plasmid was constructed to express the secretory insulin from a different promoter, to eliminate the possibility that different secretion levels may arise from a regulatory factor affecting the MF*α1* promoter in pDP314-Ins1. The new plasmid utilised the *PGK1* promoter. Product yields were less, but the relative differences in secreted product level were maintained, supporting the view that the increases were due to gains in the secretion pathway.

As an alternative approach to analysing the mutants, they were transformed with pJK1-αw, which expresses secretory wheat α-amylase. The three mutants all

behaved differently to the wild-type parent strain, and to one another, in terms of percent amylase secreted, relative secretion yield and expression level. This suggests they could be mutations in different genes.

When mated with a wild-type strain to form a diploid, the three mutants all gave rise to similar levels of secreted mini-proinsulin as their wild-type parent, showing each of the 3 mutations to be recessive.

The diploids were sporulated for further genetic analysis. Spore-derived colonies were cultured and Elisa assays for secreted mini-proinsulin undertaken with their supernatants. They showed a 2:2 segregation of the enhanced secretion phenotype, indicating that it arose from a single chromosomally-located gene in each case.

Complementation analysis, to determine whether the 3 mutations occur in different genes is underway. If so, then a strain carrying all three mutations may show a cumulative increase of >10-fold secretion yield.

2. S. cerevisiae *mutants showing enhanced secretion of* A. niger *β-galactosidase*

The β-galactosidase of *Aspergillus niger* is a large, glycosylated molecule (polypeptide M_r 110 kDa, glycosylated M_r 130 kDa) that is very inefficiently secreted by *S. cerevisiae* (<10%). This contrasts with the small, non-glycosylated mini-proinsulin (6 kDa) used above. We have shown by spheroplasting experiments that a substantial amount of β-galactosidase activity is trapped in the periplasmic space. It was therefore perceived that mutants improving secretion efficiency may affect cell wall porosity.

Mutant screens were initiated for yeast colonies containing pVK1.1 with increased halo size on X-gal plates. A number of candidates were isolated, but their screening was complicated by a variable phenotype and toxicity of the product to the cells. In selective culture, cells quickly became flocculant and died off in subculture. Colonies on selective indicator plates tended to lose their blue colour; although this was restored by a period of non-selective culturing. These factors complicated screening. However, repeated screening eventually yielded three mutants whose secreted -βgalactosidase levels were higher: #1 by 3-fold, #5 by 2-fold, and #10 by 2-fold.

In each case the total amount of β-galactosidase produced by the cell appeared to be slightly reduced (~ 30%). It was considered possible that the improved secretion yield could be due to reduced expression causing a reduced build-up in the cell and periplasmic space. To address this question, an inducible expression system was constructed and introduced into each of the mutants. The secretory β-galactosidase coding sequence was recloned into a new plasmid vector downstream of a *PGK* promoter with an androgen responsive UAS. Each host strain was also transformed with an integrative plasmid encoding the androgen receptor. In the absence of the hormone, the transformed cells could be cultured with no detectable β-galactosidase being expressed — and hence no toxic effects. Addition of testosterone to the medium resulted in induction of β-galactosidase expression. The levels of expression were about 50-fold less than from pVK1.1. Mutants #5 and #10 showed a 2- to 3-fold greater secretion yield than wt, whilst #5 showed twice as much total product as wt, and #10 a similar total to wt. The percent secreted doubled for the wt, suggesting that lowering the expression level could reduce the build-up of cell-retained product. Mutant #1 showed double the percent secreted than wt, although the same amount was secreted, yet the total yield was half that of wt. Mutant #1 may therefore exert its effect via reducing β-galac-

tosidase expression. Studies to complete the genetic characterization of these mutants are in progress.

3. Investigation of the decline in secretion products in S. cerevisiae *culture supernatants during stationary phase incubation*

Secreted mini-proinsulin was found to decline in yeast batch cultures during prolonged stationary phase incubation after the cessation of fermentative growth, independently of pH. In contrast, wheat α-amylase — a six-fold larger protein — showed a less severe decline, independent of pH, suggesting that the small size of mini-proinsulin made it particularly susceptible to the apparent degradative loss.

No link could be made with presence or absence of mutations in vacuolar protease genes or cell mating type. Thus, neither leaky secretion of vacuolar proteases nor rare cell lysis seemed to be the cause.

Incubation of 1000-fold concentrated yeast culture supernatants (pH neutralised) with human serum albumin, or bovine, resulted in the production of a proteolytic fragment, similar in size to that reported to occur when HSA is secreted.

In contrast, secreted *A. niger* β-galactosidase increased during prolonged stationary phase incubation rather than decreased. Likewise, *A. niger* amyloglucosidase added exogenously to yeast cultures showed no loss during prolonged incubation. Thus, neither of these proteins were sensitive, but both are large and glycosylated. The possibility that glycosylation was protecting them from degradation was investigated. The β-galactosidase expressed in an *mnn1 mnn9* mutant strain, which does not add outer chain glycosylation, showed an early high level of product in the culture medium, that delined to a low, constant level rather than increasing, as with the wt stain. Additionally, endoH treated and refolded (deglycosylated) invertase was degraded when incubated with concentrated culture supernatant, whereas the untreated protein was not. Thus, glycosylation appeared to protect secretory proteins from proteolysis.

Activity gels with concentrated culture supernatants show (low-level) proteolytic activity with a general substrate and with a specific substrate for trypsin-like activity.

When cells were treated with brefeldin (to block secretion), the cells and supernatants separated, with the cells being resuspended in fresh medium containing brefeldin, much stronger insulin degradative activity was associated with the cell fraction than the supernatant fraction. This suggests that the degradative activity is mainly cell-associated, with that in the medium probably due to leaching.

4. Evaluation of yeast secretion systems for the production of trout somatotropin

An expression cassette containing the trout somatotropin cDNA sequence under the control of a strong constitutive promoter (GAPDHp) and the prepro signal sequence of the α mating factor was introduced in different yeast strains on a high copy vector. Different media at different pH were tested. Secretion of the somatotropin was never observed. Instead it accumulated intracellularly in an insoluble form.

As the expression of this protein may be toxic to the host, other expression cassettes were assembled using regulated promoters. When the expression was driven by a *ADH2-GAPDH* hybrid promoter and the pre signal of the *Kluyveromyces* killer toxin, most recombinant somatotropin accumulated intracellularly in an insoluble form. Modifications of the growth medium and of inductive conditions failed to improve these results. In another construct, the *ADH2-GAPDH* promoter was

replaced by the *HSP12* promoter. Different strains and different growth conditions were tested. In all strains most of the recombinant protein was found in the intracellular insoluble fraction at levels up to 100 mg/l. An intracellular soluble fraction was also observed (max. 4 mg/l). In some strains a low level of secretion was obtained. At 30°C up to 80 µg/l and at 20°C up to 300 µg/l were secreted. Good results were obtained in a radioimmuno assay and in a radioreceptor assay indicating that the protein was most probably biologically active.

Attempts were made to mutagenise yeast cells to improve their capacity to secrete the trout somatotropin. After different mutagenising treatments, yeast cells were screened in a plate footprint immunoassay. No clone showing an improved secretion phenotype could be isolated.

Finally, to improve the secretion of this somatotropin, new expression cassettes were constructed to produce translational fusions with the wheat α-amylase. This protein is efficiently secreted by *S. cerevisiae* and could force the secretion of a fused protein. This work is still in progress.

5. Lactose fermentation by brewing yeast secreting A. niger *β-galactosidase*

S. cerevisiae does not ferment lactose, although it will do so if transformed to express the secretory β-galactosidase of *A. niger*. Thus, a brewing yeast containing pVK1.1 will ferment lactose and produce ethanol. When whey permeate is used (a waste product of the dairy industry) with a content of 10% lactose, at the end of fermentation 5% alcohol is produced, with no lactose remaining. This provides the possibility of new product development, as whey permeate is commonly used as a food additive.
A brewing strain containing an integrated form of the β-galactosidase gene showed evidence of instability after a period of culture. Isolates failing to ferment lactose were found to have lost the lactase gene; others showed different levels of activity. A variant was selected that
showed increased activity. In trials this also resulted in a much more rapid decrease in specific activity of the wort during fermentation.

C. Functions governing cell wall structure and integrity

1. Cell wall structure

Three exo-1,3-β-glucanase-encoding genes (*EXG1, EXG2* and *SSG1*) in *S. cerevisiae* were cloned and characterized. *EXG1* codes for a 448-aa polypeptide whose differential glycosylation at the two potential N-glycosylation sites accounts for the two main extracellular exo-1,3-β-glucanases (Exo I and Exo II) detected in culture supernatants. Amino-terminal sequence analysis of the secreted proteins reveals the existence of an extra 40-aa peptide in the precursor protein containing a Lys-Arg peptidase processing site (KEX2) at the junction of the mature, extracellular forms. Site-directed mutagenesis of the two N-glycosylation sites indicates that Exo I originates from Exo II by selective elongation of the oligosaccharide chain bound to the second sequon.

Sequencing analysis of the *EXG2* gene shows a 1686-bp ORF encoding a polypeptide of 562-aa, which contains 14 potential N-glycosylation sites. The mature product (Exo III) is a high molecular weight protein (200 kDa) exhibiting a high carbohydrate content. Contrary to the *EXG1*-encoded glucanase, only about 5% of the Exo III is detected in the culture supernatant. Comparison of the predicted amino acid sequence of this exo-glucanse with those of other reported yeast cell-wall-associated proteins reveals a C-terminal structure characteristic of polypep-

tides attached to the plasma membrane through a glycophophatidylinositol (GPI) anchor. Deletion of the hydrophobic C-terminus or point modificatioon of the putative GPI-attachment site (KNNA) releases the protein to the surrounding medium, thus pointing to a direct involvement of the carbxy-end of the primary translation product in the correct final location of the mature wild-type Exo III glucanase.

1,3-β-glucanase-encoding genes from yeasts

	*% similarity**	*nucleotides*	*aminoacids*	*molecular mass*	*glyco-sylation sites*
EXG1 S. cerevisiae	100.0	1344	448	51307	2
EXG2 S. cerevisiae	33.0	1686	562	63530	14
SSG1 S. cerevisiae	75.0	1335	445	51790	1
EXG K. lactis	69.4	1287	429	49801	0
EXG S. occidentalis	66.5	1275	425	49130	1
EXG H. polymorpha	56.5	1305	435	49255	2
EXG Y. lipolytica	60.0	1263	421	48093	0
ENG1 S. cerevisiae	—	3351	1117	121077	10

The *SSG1* gene specifies a 445-aa sporulation-specific exo-1,3- β-glucanase, which can be detected only in sporulating diploids (a/α) , but does not appear in vegetatively growing cells or in non-sporulating diploids (α/α) when incubated under nitrogen starvation conditions. The meiotic course of *SSG1* induction indicates that the gene is transcribed only in the latter stages of the process, beginning at the time of meiosis and reaching a maximum during spore formation. Homozygous ssg1/ssg1 mutant diploidsare able to complete sporulation, although with significant delay in appearance of mature asci. Deletion ans substitution studies at the 5 -noncoding region identifies a 99-bp upstream activiation sequence (UAS, -1096 to -997), which is essential for sporulation-regulated expression of the gene.

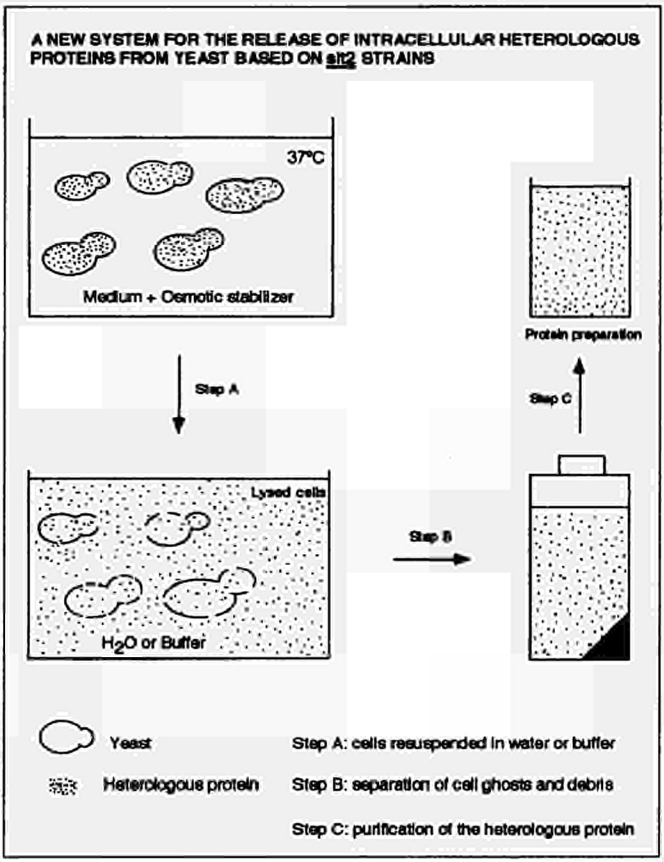
Comparison of the predicted amino acid sequences of the *EXG1, EXG2* and *SSG1* products revealed five highly conservedregions, which are located in the same relative positions in the three polypeptides and may be essential for β-glucanase function. removal of the histidine and tyrosine residues at one of these regions (DHHHY) abolishes β-glucanase activity. Replacement of the second histidine or the tyrosine results in dramatic loss of avtivity, suggesting that these residues are fundamental for catalysis and/or binding to the substrate.

Taking advantage of the high degree of similarity exhibited by the exo-glucanases from *S. cerevisiae*, we have carried out the cloning of exo-glucanase genes from non-conventional yeasts (*Scwanniomyces occidentalis, Klyuveromyces lactis, Hansenula polymorpha* and *Yarrowia lypolytica*) through PCR-amplification. Comparison of the amino acid sequences of the predicted primary translation products encoded by these exo-glucanase genes with those encoded by the previously characterized *EXG1, EXG2* and *SSG1* genes from *S. cerevisiae* revealed that members of this protein family are highly conserved throughout evolution.

The molecular cloning of a gene (*ENG1*) encoding an endo-1,3-β-glucanse from *S. cerevisiae* was achieved. *ENG1* encodes a 1117-aa polypeptide which is efficiently secreted to the periplasmic space and then released to the culture medium.

Osmotically stable phenotype

The gene *SLT2* codes for a serine/threonine MAP kinase that is essential for the development of an osmotically stable cell wall. This kinase participates in a novel cascade of phosphorylation reactions starting with the mammalian protein kinase



homolog Pkc1. The functionality of this gene was defined through the development of null mutants, point mutants affected in essential regions and regulation of the expression of the gene. The phenotype of these mutants usually led to thermosensitive osmotic-remedial autolysis of the cells. A thorough characterization of the functionality of gene *SLT2* was achieved through the definition of the autolytic osmotic-remedial phenotypes of null and point mutants. The conditions for protein release from *slt2* mutants were clearly established in fermenter cultures (10 l to 15 l scale) both in batch and continuous cultures. Protease-deficient *slt2* strains were also developed that could be used for the release of heterologous proteins (bacterial cloramphenicol-acetyl-transferase was used as the model system). The extent of release was of the order of 70% of the total intracellular heterologous protein. It was achieved both by switching temperature of growth to non-permissive conditions and by osmotic shock of cells previously incubated at the non-permissive temperature with the protection of 0.5M sorbitol.

Another model of autolytic mutants, namely those affected in gene *LYT1*, that expressed a phenotype non-osmotic remediable was also characterized. The cells lysed upon expression of a mutation that was characterized as cdc-like and a suppression of the mutation was achieved with the previously described sporulation gene *SPO12*.

Other significant achievements were the isolation and characterization of a set of autolytic osmotic-remediable mutants from diploid yeast and the development of a model system of gene reporter in yeast based on the expression of exo-1,3-β-glucanase, that can be followed by means of Flow Cytometry.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The *HSP12* promoter was shown to be particularly good for recombinant trout somatotropin expression, and the strongly-induced, thiamine regulated promoter has good prospects for biotech applications. Enhanced secretion mutants obtained provide an alternative approach to analysing the secretion pathway, which is known to be complex, and have direct potential for biotech exploitation. There is commercial interest in novel brews utilizing fermented whey. A solid basis for characterization of functions governing yeast cell integrity has been established. Autolytic strains, that are also protease deficient, have been developed for the release of heterologous intracellularly-expressed proteins and fermenter systems devised, although the systems are still under development.

MAJOR COOPERATIVE LINKS

Regular biannual meetings. Exchange of information, materials and staff. Major interactions: fermentation (B, E, G); secretion (A, B, C, E, F, G); expression (D, G, B); cell wall manipulation (E, F, A, B).

PUBLICATIONS

Joint publications

Cid. V.J., Duran, A., del Rey, F., Nombela, C. and Sanchez, M. (1994). The molecular basis of cell integrity and morphogenesis in *Saccharomyces cerevisiae*. Submitted to Microbiological Reviews. (E, F)

Gill, B., Pioli, D. and Hadfield, C. Recombinant polypeptide production by supersecreting mutants. (in preparation). (A, B)

Individual publications

Correa, J., Vazquez de Aldana, C.R., San Segundo, P. and F. del Rey (1992). Genetic mapping of 1,3-β-glucanase-encoding genes in S. cerevisiae. Current Genetics **22**: 283-288.

San Segundo, P., Correa, J., Vazquez de Aldana, C.R. and F. del Rey (1993). SSG1, a gene encoding a sporulation-specific 1,3-β-glucanase in S. cerevisiae. Journal of Bacteriology **175**: 3823-3837.

de la Fuente, J.M., Vazquez, A., Gonzalez, M.M., Sanchez, M. Molina, M. and Nombela, C. (1993). Expression of mutations and protein release by yeast conditional autolytic mutants in batch and continuous cultures. Applied Microbiology and Biotechnology **38**, 763-769

Hadfield, C., Raina, K.K., Menon, K.S. and Mount, R.C. (1993). The expression and performance of cloned genes in yeasts. Mycological Research **97**, 897-944.

Martin, H., Arroyo, J., Sanchez, M., Molina, M. and Nombela, C. (1993). Activity of the yeast MAP kinase homolog Slt2 is critically required for cell integrity at 37 C. Molecular and General Genetics **241**, 177-184.

Molero, G., Yuste-Rojas, M., Montesi, A., Vazquez, A., Nombela, C. and Sanchez, M. (1993). A cdc-like autolytic *Saccharomyces cerevisiae* mutant altered in budding site selection is complemented by SPO12, a sporulation gene. Journal of Bacteriology **175**, 6562-6570.

Cid, V.J., Sanchez, M. and Nombela, C. (1994). Characterization of thermosensitive autolytic mutants from diploid *Saccharomyces cerevisiae*. Microbiology **140**, 559-568.
Hadfield, C. (1994). Construction of cloning and expression vectors. In Molecular Biology of Yeast: A practical Approach (ed. J.R. Johnston), pp. 17-48. IRL Press: Oxford, UK.
Cid, V.J., Alvarez, A.M., Santos, A.I., Nombela, C. and Sanchez, M. (1994). Yeast exo-b-glucanases can be used as efficient and readiliy detectable reporter genes in Saccharomyces cerevsisiae. Yeast. (in press).
Alvarez, P., Sampedro, M., Molina, M. and Nombela, C. (1994). A new system for the release of heterologous proteins from yeast based on mutant strains deficient in cell integrity. (in preparation).
Gill, B., Mount, R.C., Dawson, D. and Hadfield, C. (1994). Yeast mutants isolated for enhanced secretion of mini-proinsulin. (in preparation).
Gill, B. and Hadfield, C. (1994). The isolation and genetic characterization of yeast mutants showing enhanced secretion of *A. niger* β-galactosidase. (in preparation).
Turver, C., Gill, B., Dawson, D. and Hadfield, C. (1994). Analysis of degradative activity that can reduce the yield of secreted proteins produced by *S. cerevisiae* cultures. (in preparation).
Praekelt, U., Hather, R.J., Byrne, K. and Meacock, P.A.. Characterisation of a set of genes induced when molasses-grown yeast cultures enter stationary phase. (in preparation)
Hather, R.J., Praekelt, U.M. and Meacock, P.A.. A new thiamine-regulated gene of *Saccharomyces cerevisiae*, THI5. (in preparation).
Hather R.J. & P.A. Meacock. The YCR107W gene of *S. cerevisiae* chromosome III is a member of a multi-gene family that includes pseudogenes. (in preparation)
Basco, R.D., Hernandez, L.M., Muñoz, M.D., Olivero, I., Andaluz, E., del Rey, F. and G. Larriba. 1994. 'Selective elongation of the oligosaccharide attached to the second potential glycosylation site of yeast exoglucanase: effects on the activity and properties of the enzyme'. (submitted).

Wide domain control of primary and secondary metabolism in Aspergilli (BIOT CT-900169)

COORDINATOR:
1. J. VISSER, MGIM, Wageningen Agric. University, Wageningen, NL
PARTICIPANTS:
2. H.N. ARST (Jr), Royal Postgrad. Med. School, London, GB
3. C. SCAZZOCCHIO / B. FELENBOK, Inst. Génét. et Microbiol., Orsay, F
4. M.A. PEÑALVA, Centro de Investigaciones Biologicas (CSIC), Madrid, E
5. F. NALIN, Lyven, Cagny, F

BACKGROUND INFORMATION

The filamantous fungus *Aspergillus nidulans* has become during the last decades the most advanced model system to study regulation of fungal gene expression, in connection with fungal physiology. This project was launched to enlarge our understanding of basic regulatory mechanisms and to explore whether this knowledge could be expanded to filamentous fungi of industrial importance *(Aspergillus niger, Penicillium chrysogenum).* Two wide domain regulatory systems, carbon catabolite repression and pH regulation, were of particular interest as these are also major systems involved in the regulation of industrially important products (antibiotics, enzymes, primary metabolites).

OBJECTIVES AND PRIMARY APPROACHES

(1) Defining sites of action of the CREA product for *alc* and *prn* genes (F).

(2) Characterization at the molecular level of the relationships between induction and carbon catabolite repression (F).

(3) Defining *cre*A target systems in *A. nidulans* and *A. niger* (NL).

(4) Metabolic analysis of mixed carbon source utilization in *A. nidulans* and *A. niger* (wild type, mutants) (NL).

(5) Obtaining and physiologically characterizing a null *pac*C allele in *A. nidulans* (GB).

(6) Selection and sequencing of a phenotypically wider range of *pac*C mutant alleles in *A. nidulans* to understand more about *pac*C structure and function (GB).

(7) Further characterization of target genes for use in studying *pac*C regulation (GB, NL)

(8) Studies on the effects of *pac*C and *pal* mutations on the repression of the IPNS (= *ipn* A) gene (E, GB).

(9) Sequencing of the *A. niger pac*C gene and characterizing *A. niger pac*C mutants.

(10) Continuation of metabolic and pH analysis of *A. nidulans* wild type and mutants including ^{13}C- and ^{31}P-NMR spectroscopy.

RESULTS AND DISCUSSION

A. Carbon catabolite repression

1. Defining sites of action of CRE A for the alc *and* prn *regulons and characterizati on of the relationship between induction and carbon catabolite repression* (Orsay)

The characterization at a molecular level of the relationship between specific induction and carbon catabolite on the two regulons, *alc* and *prn* in *Aspergillus nidulans* has considerably progressed. The results scheduled during the period 24-36 month have all been obtained.

— Production in *E. coli* of regulatory proteins using fusions with the glutathione-S-transferase (GST). The DNA binding domains of the three regulatory proteins from Aspergillus nidulans CREA, ALCR and PRNA are available as fusion proteins isolated in *E. coli*: GST::ALCR (7-60), GST::PRNA (18-272), and GST::CREA (35-247). Gel band shift experiments and footprinting studies using various chemicals and protection to DNaseI were performed with these fusion proteins.

— Identification of induction sites.

a) *In the ethanol regulon.* The work on the ethanol regulon has been completed by a functional analysis of the ALCR binding sites in the *alc*R and *alc*A promoters.
Both promoters contain invert and direct repeat targets with the same consensus core. However, the mechanism of transcriptional activation is different in both genes. In the *alc*R promoter, the invert repeat target is sufficient to drive a full transcriptional activation of the autoregulated *alc*R gene. In the *alc*A promoter, the three ALCR binding sites are necessary for full induction. Deletion analysis and site directed mutagenesis of the three binding sites have shown that the ALCR transcriptional activation is synergistic by a cooperative binding mechanism of ALCR on its cognate targets.

b) *In the proline cluster.* The PRNA binding sequences (CCGG10-15CCGGNCCGG) have been identified by band shift and foot printing analysis in the *prnB-D* intergenic region and in the *prn*C promoter. No PRNA sites have been found for prnX, which suggests an at-distance effect from the *prn*D promoter. Transcriptional studies of strains carrying different deletions have allowed to establish the functional relevance of the PRNA sites identified *in vitro*.

— Identification of carbon catabolite represssion sites.

a) *In the ethanol regulon.* Several CREA binding sites with the consensus binding 5'G/CYGGPuG3' have been localized by gel retardation assays and foot printing analysis. In the *alc*R promoter 8 binding sites exist, in the *alc*A promoter 7 have been localized.
By deletion analysis and site directed mutagenesis, one CREA binding site in the *alc*R promoter has been shown to be functional. The disruption of this CREA binding site resulted in a phenotype which was highly derepressed (80%) and superinduced (100 fold).
This result was the first evidence of a direct relationship between induction and carbon catabolite repression (see below).

b) *In the proline cluster.* A thorough molecular analysis of all CREA in vitro binding sites in the *prnB-D* intergenic region has been performed. Efficient repression needs the integrity of two of these sites which are adjacent and divergently oriented. The minimal requirements for CREA binding and the role of the context of CREA sites have been established by band-shift

analysis. Whereas the binding ability of SYGGGG type sites is absolute, and context independent, the presence of an adenine in the fifth position demands a certain sequence context to bind CREA.

— Relationships between specific induction and carbon catabolite repression.
The ethanol regulon is subject to specific induction and carbon catabolite repression. The proline cluster is subject to these forms of control but also to nitrogen metabolite repression. In the ethanol regulon, the CREA repressor is a major switch controlling the expression of the ethanol genes under all physiological conditions, derepressed and repressed.
An important level of control is the almost full repression of the specific positive acting gene *alc*R and the direct repression of the structural genes. This repression occurs by binding of CREA at its cognate GC rich targets. A direct competition exists between the transactivator ALCR and the repressor CREA for the same promoter regions: the binding of one regulator preventing the binding of the second one. In the proline cluster, repression of the specific regulatory gene *prnA* is marginal. Efficient repression occurs via a cis-acting region containing the CREA binding sites. It was suggested that CREA acts by preventing the activity of a different and possibly general transcription factor.

2. *Defining* creA *target systems in* A. nidulans and A. niger *and metabolic analysis of mixed carbon sources* (Wageningen)

A. nidulans strains carrying *cre*A mutant alleles were used to investigate the role of *cre*A on the inducibility and expression levels of several extracellular enzymes. The expression of two *A. nidulans* polypeptides which are induced by pectin and which cross-react with monoclonal antibodies raised against pectin lyase A (PLA) from *A. niger* has been analyzed. Expression has been observed to be repressed in glucose medium. In the *creA^{d}30 background a much higher inducibility of these putative pectin lyases was found than in the wild type. Similarly, two araban degrading enzymes* (α-L-arabinofuranosidase and endo-arabinase) were found to be highly overexpressed in *A. nidulans cre*A^{d}30 and *cre*A^{d}4 under inducing conditions. L-arabinose catabolism itself is also under control of *cre*A and therefore affects arabinase biosynthesis in a more indirect way via the intracellular levels of the pathway intermediate L-arabitol. This compound has been shown to be the inducer for arabinase biosynthesis.
We have succeeded in generating two *A. niger* mutants (*cre*A1 and *cre*A2) which have been proved to be *cre*A$^-$ by virtue of their complementation by the cloned *cre*A$_{nig}$ gene. Preliminary studies of the effects of the *cre*A alleles on the expression of the *A. niger pel*A gene and the arabinase genes *abf*A, *abf*B and *abn*A have indicated that the two *cre*A alleles give rise to different effects dependent on the particular target gene analysed. These data indicate a role for *cre*A in the expression of pectin lyase and arabinase genes both in *A. nidulans* and *A. niger*.
Recombinant plasmids carrying a fusion gene comprising the *A. niger* zinc finger (Zf) coding region and glutathione-S-transferase (GST) have been constructed using PCR techniques.

B. pH Regulation

1. *Selection, sequencing and physiological characterization of a range of pacC mutant alleles* (London)

A null allele was constructed using homologous recombination with a clone deleted for the entire coding region, which was replaced by the *pyr*4 gene of Neurospora crassa, allowing complementation of a *pyr*G auxotrophy of *A. nidulans*. The

phenotype of this allele resembles that of the pal mutations in mimicking the effects of growth at acid pH but in addition it reduces growth and conidiation generally.
A number of new mutations have been selected and sequenced, including missense, frame-shift and chain termination mutations. Frame-shift and chain termination mutations in the C-terminal region in general mimic growth at alkaline pH whereas truncations further upstream mimic at least in part the effects of acid growth conditions. As yet, apart from the null allele, no mutations have been obtained in the DNA binding region. Very recently, however, a new phenotypic class of mutations, mimicking strongly growth at acid pH for some activities and growth at alkaline pH for others, has been obtained. As these mutations have been obtained so recently, the sequence changes have not yet been determined.
The regulation of the *pac*C mRNA has been studied. It is highest at alkaline growth pH and in mutants mimicking alkaline growth pII and lowest at acid pH and in mutants mimicking acid growth pH.
Further sequence data has been obtained for the acid phosphatase structural gene *pac*A and for the promoter region of the alkaline protease structural gene *prt*A.
The pattern of pH regulation of the *pac*C message, the sequence changes in *pac*C classical mutants and the phenotype of a *pac*C null mutation all indicate that revision of the pH regulation model is necessary. That the PacC protein is a transcriptional factor seems beyond doubt. The roles of the *pal* genes in synthesising an effector for the PacC protein continues to be the most likely possibility for these genes. However, it is now clear that PacC activates transcription of genes preferentially expressed at alkaline pH and prevents transcription of genes expressed at acid pH. The effector produced by the *pal* pathway is necessary for both the negative and positive roles of PacC.

2. pH and carbon regulation of the isopenicillin N synthetase gene (Madrid, London)

In close collaboration the two laboratories have shown that the pH regulatory system is the major determinant in expression of the *ipn*A gene. External alkaline pH substantially elevates *ipn*A transcript levels in the presence of a repressing C source, as shown by transcript analysis from mycelia grown in sucrose broth buffered at alkaline pH, or grown at acidic pH and subsequently transferred to alkaline pH media for a short period of time. All tested mutations in *pac*C which mimic the effects of growth at alkaline pH also result in substantial elevation in *ipn*A transcript levels and also elevation of penicillin levels.
Under the growth conditions stated above, it was concluded that (i) *ipn*A transcription is under the direct or indirect control of PacC; (ii) penicillin biosynthesis is under the control of the pH regulatory system and (iii) pH regulation of *ipn*A transcription overrides C regulation.
All C sources found to repress *ipn*A acidify the external pH, whereas all derepressing C sources result in external alkalinisation. This raised the possibility that C regulation could be mediated through pH regulation. However, external acidic pH does not prevent C derepression caused by utilization of two different derepressing C sources (neither upon continuous growth at acidic pH nor in transfer experiments). Moreover, loss of function mutations in *pal*A, *pal*B or *pal*F (which mimic the effects of growth at acidic pH) do not prevent C derepression. Therefore, the conclusion is that C regulation does not act through pH regulation but rather that both types of regulation are mechanistically independent. This is in agreement with C regulation acting through a negative mechanism, whereas pH regulation of *ipn*A transcription requires the positive action of PacC. Moreover, the decrease in *ipn*A transcript levels caused by a repressing C source is substan-

tially more pronounced in a *pal* background. This additivity further supports the existence of independent regulatory mechanisms. Finally, despite of the fact that *pal*⁻ mutants show essentially derepressed levels of the *ipn*A transcript when grown in a derepressing C source, they do not produce penicillins, indicating that, in addition to *ipn*A, there must be a second target for pH regulation in the penicillin pathway.

3. *The* A. niger pac*C gene; characterization of target genes to study* pac*C regulation* (Wageningen)

Comparison of preliminary DNA sequence data to the *A. nidulans pac*C sequence showed the presence of similarities to the translation product of the *A. nidulans* pacC gene. The 5.4kb *A. niger* fragment has now been sequenced in its entirety. Located in this fragment is a single open reading frame which encodes a triple zinc-finger (Zf) protein of 667 amino acids and projected molecular mass 70.5kDa. Although there is considerable similarity between the primary structures of the PACC proteins of *A. niger* and *A. nidulans*, some variations occur. 2.5 kb of upstream sequence and 0.5kb of 3' terminal sequence have also been determined. The pattern of acid phosphatase expression in *A. nidulans pac*C11 transformants carrying either the *A. nidulans* or *A. niger pac*C genes using non-denaturing polyacrylamide gel electrophoresis has been analysed. Production of acid phosphatase at pH 4.5 is restored in transformants carrying either *pac*C gene compared to both the *pac*C11 mutant and transformants carrying only a recombinant argB gene.
Using the sequence data of the *A. niger pac*C gene a disruption mutant (pac2) has been generated by molecular means.
The consequences of disruption of the *pac*C gene with regard to the expression of pH regulated genes (specifically phosphatases and proteases) are currently being studied. Another *pac*C target is the *A. nidulans pep*A gene encoding an acid protease. This gene has now been completely sequenced.
Analysis of the pattern of transcription of the *A. niger pac*C gene has been undertaken in both wild type *A. niger* and the *pac*2 mutant. Transfer experiments indicate that the *pac*C gene is transcribed under alkaline conditions (pH 8) but not at acidic pH (4.5); a reduced signal has been observed at pH 6.5.

4. *pH analysis of* A. nidulans, *wild type and mutant strains altered in their pH regulation* (Wageningen)

The homoeostasis of intracellular steady state pH in *A. nidulans* wild type and *pal A1, B7, C4* and *pac*C11 mutants under different external pH conditions and with variable aeration levels was studied using ^{31}P NMR. External aeration is the main factor determining the cytoplasmic pH levels in all *Aspergillus* germlings (grown for 5 ½ hrs). Strong aeration (± 80% O_2) results in relative alkaline pH_i values between 7.5 and 7.75, whereas less aeration between 7 and 30% O2 leads to more stable conditions and to lower values of pH_i 7.2 to 7.5. Under standard aeration conditions the steady state pH_i in *A. nidulans* wild type is maintained nearly constant between pH_i 7.32 and 7.43 when the external pH_{ex} is varied in the range from 4.6 to 7.65 pH units. Thus pH homoeostasis results in an intracellular pH 1 to 3 pH units more alkaline than the external pH when the latter is between pH_{ex} 4 and 6.5. Conversely, in external alkaline medium (pH_{ex} > 7.3) the pH_i is kept more acidic inside the cells than the external medium. In contrast to the wild type, *A. nidulans pal A1, B7* and *C4* pH mutants regulate their intracellular steady state pH on a more alkaline level between pH_i 7.4 and 7.7 when the external pH_{ex} is varied between 4.6 and 6.3 pH units. In medium above pH_{ex} 6.7 the *palA1* mutant

keeps its intracellular pH_i also more alkaline than the external medium, which is in direct contrast to the wild type which is able to maintain an internal pH_i more acidic than the external pH_{ex} under alkaline conditions. This lack of pH_i regulation in the *pal* mutants at high external pH's might be a reason for their inability to grow at external alkaline pH as described. Additionally all *pal* mutants lost intracellular phosphate to the surrounding medium during an 2 ½ hrs lasting *in vivo* measurement. Data on the intracellular pH in *pac*C 11 mutants are only preliminary and show a tendency that the intracellular pH_i seems to be more acidic as in wildtype under these conditions when the external pH_{ex} of the medium is 6.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

(1) After the identification of ALCR, PRNA and CREA binding sites the mechanism of carbon catabolite repression in relation to specific induction has been worked out for the very first time in a hyphal fungus using the *alc* and *prn* systems as a model.

(2) Carbon catabolite repression of the biosynthesis of several inducible extracellular enzyme systems (pectinases, arabinases) has been shown to be mediated through *cre*A.
Repression occurs both at the level of inducer formation and extracellular enzyme biosynthesis. This provides the basis for a general regulatory model for the biosynthesis of extracellular polysaccharide degrading enzymes.

(3) Considerable progress has been made in the molecular analysis of *pac*C and its regulation necessitating modification of the model for pH regulation.

(4) The *ipn*A gene is the first well characterised structural gene found to be under positive control by *pac*C. The demonstration that a structural gene for penicillin biosynthesis is subject to direct control by the pH regulatory system is a major advance in elucidating the (largely obscure) regulation of secondary metabolism in fungi.

(5) In the penicillin pathway there are at least two targets for pH regulation. pH and carbon regulation of penicillin biosynthesis normally act in concert and are mechanistically independent. In the case of *ipn*A transcription, pH regulation was shown to override carbon catabolite repression.

(6) Using *in vivo* ^{31}P-NMR spectroscopy reliable values for the intracellular steady-state pH of well-aerated *A. nidulans* germlings could be determined as a function of the external pH. *A. nidulans pal* mutants were found to regulate their pH_i at a more alkaline level whereas the mutant *pac*C11 tends to be more acidic.

Wider considerations of the scientific results are:

(1) The effective competition between specific induction and carbon catabolite repression provides a new tool to obtain superinduction of the *alc*R gene by disrupting CREA targets. Since *alc*A transcription is related to that of *alc*R this provides an improved *alc*R-*alc*A system for expression of heterologous proteins.
Physiological studies using *A. nidulans* and *A. niger cre*A mutants also revealed such a mechanism of superinduction for several extracellular enzymes.

(2) *A. nidulans* continues to be the organism where pH regulation is best characterised and understood. In addition to its scientific novelty pH regulation is of

very considerable practical importance as illustrated by its role in penicillin biosynthesis and in the synthesis of extracellular enzymes such as proteases.

(3) Since *A. niger* was found to have a very similar pH regulatory system which can be functionally exchanged, this illustrates the potential of *A. nidulans* as a model to develop knowledge in industrial fungi. Similarly, since *A. nidulans* strains carrying gain-of-function mutations in *pac*C are penicillin overproducers in the presence of a repressing carbon source this suggests an obvious approach to generate *Penicillium chrysogenum* overproducers through reverse genetics.

(4) The overriding role of pH regulation in the expression of the *ipnA* gene strongly suggests that alkaline pH represents a physiological signal which triggers penicillin biosynthesis. Limitation for carbon most likely results in the catabolism of amino acids as C source with the consequent release of ammonia which would in turn alkalinise external pH. This might represent a signal in response to which the fungus releases antibiotics to harvest C compounds from sensitive bacteria. In addition, it is hard to imagine that bacteria, which usually grow at very narrow (neutral) pH range, may seriously compete with fungi under acidic external conditions.

(5) The information obtained by ^{31}P-NMR spectroscopy and spectrofluorimetry about the cytoplasmic pH_i and the vacuolar pH of *A. nidulans* germlings is, together with external pH measurements, important to understand the role of these compartments and the various pH gradients which exist in pH regulation, in transport processes and in the regulation of metabolic fluxes.

Industrial participation was realized in the following way:
Representatives of Lyven (Cagny, France), Antibioticos (Léon, Spain) and Ciba Geigy AG (Basel, Switzerland) have participated in the periodic meetings whereas the latter two companies have assisted in organizing the 3rd annual project meeting. Contacts between M.A. Peñalva (Madrid) and Antibioticos, between H.N. Arst Jr and Ciba Geigy and also between J. Visser and Ciba Geigy have contributed to an active industrial participation which is also reflected by joint publications.

MAJOR COOPERATIVE LINKS

a) Meetings

M.A. Peñalva and C. Scazzocchio have succesfully organised the EMBO workshop on Molecular Biology of Filamentous Fungi (September 18-22, 1993) which is a major meeting in the field. This occasion was used to organize a separate project meeting.
A combined BRIDGE-BIOTECH project meeting took place in Wageningen (February 26-27, 1994).
J. Visser, J.P.T.W. van den Hombergh and H. Panneman assisted D. von Wettstein (Copenhagen) in organizing and teaching the EMBO Advanced Course 'Production of Heterologous Proteins in Yeast and Filamentous Fungi(August 20-30, 1993).

b) Staff exchange

H.N. Arst Jr visited the Orsay laboratory and examined two doctoral theses in December 1993. In addition he and J. Tilburn visited the Madrid laboratory for collaborative work and to participate in the writing of a joint publication about the role of *pac*C in regulating *ipn*A transcription.

J.P.T.W. van den Hombergh spent a 2 months period in Basel with F. Buxton of Ciba Geigy.

c) Other cooperative actions.

The participating laboratories were all in frequent telephone, telefax and E-mail contact concerning collaborative projects and to exchange scientific ideas. Contacts of M.A. Peñalva with the Paris laboratory (C. Scazzocchio, B. Cubero) have for instance been essential for the timely publication of the study on the *in vitro* CREA binding sites present in the *ipn*A upstream region. Contacts between the Wageningen and London laboratories have resulted in a manuscript which has been submitted whereas two other joint manuscripts are in preparation. A large number of *A. nidulans* strains were supplied by the London laboratory to the Madrid, Orsay and Wageningen laboratories. Further exchange of strains and genes between the four participating laboratories has been continuous and fluent.

d) European dimension

The collaborative attitude of all partners involved, their mutual understanding and their shared interest in basic problems in fungal gene regulation and physiology, has established a scientifically stimulating and productive climate. Besides the important effect the project has in removing borders, it has also stimulated research at the national level. This project is experienced by all participants to be important as a nucleus to establish a more expanded european network in fungal biotechnology in the near future.

PUBLICATIONS

Joint publications

Espeso, E.A., Tilburn, J., Arst, H.N., Jr. and Peñalva, M.A. (1993) pH regulation is a major determinant in expression of a fungal penicillin biosynthetic gene *EMBO J.* **12**, 3947-3956.
Hombergh van den, J.P.T.W., Jarai, G., Buxton, F.P. and Visser, J. Molecular cloning, characterization and expression of pepF, a gene encoding a serine carboxy peptidase from *Aspergillus niger* Gene (in press).
Jarai, G., van den Hombergh, J.P.T.W. and Buxton, F.P. (1994), Cloning and characterization of the pepE gene of *Aspergillus niger* encoding a new aspartic protease and regulation of the *pep*E and *pep*C genes. Gene **145**, 171-178.
Macrac, W.D., Buxton, F.P., Sibley, S., Garven, S., Gwynne, D.I. Arst, H.N., Jr. and Davies, R.W. (1993) Charcaterization of an *Aspergillus nidulans* genomic DNA fragment conferring phosphate-non-repressible acid phosphatase activity. Gene **130**, 247-251.
Ramón, D., van der Veen, P. and Visser, J. (1993) Arabinan degrading enzymes from *Aspergillus nidulans*: Induction and purification FEMS Microbiol. Lett. **113**, 15-22.
Tilburn, Espeso, E.A., Orejas, M., Peñalva, M.A., J., Sarkar, S., Widdick, D.A., Bignell, E., Mungroo, J. and Arst, H.N., Jr. Molecular characterisation of the *pac*C gene mediating pH regulation of gene expression in Aspergillus nidulans (in preparation).
Veen van der, P., Flipphi, M.J.A., Arst, H.N. Jr. and Visser, J. Extracellular arabinases in *Aspergillus nidulans*: the effect of different cre mutations on enzyme levels. Arch. Microbiol. (in press).

Individual publications

Arst, H.N., Jr. (1994). Regulation of Gene Expression by Oxygen, Phosphorus and pH. In: Martinelli, S.D. and Kinghorn, J.R. (eds) *Aspergillus*: 50 years on, Amsterdam, Elsevier, pp. 369-380.
Cubero, B., and Scazzocchio, C. Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans. EMBO J.*, **13**, 405-415 (1994).

Espeso, E.A. and Peñalva, M.A. (1994). *In vitro* binding of the two finger repressor CreA to several consensus and non-consensus sites at the *ipnA* upstream region is context dependent. FEBS Letters, **342**, 43-48 (1994)
Felenbok, B., and Sealy-Lewis, H.M. (1994) Alcohol metabolism. In: *Aspergillus*: 50 years, edited by Martinelli, S. and Kinghorn, J.R., Elsevier Scientific, Amsterdam, pp. 141-180.
Felenbok, B., and Kelly, J. The regulation of carbon metabolism in filamentous fungi. In: The Mycota, Vol IV, edited by Marzluf. G. and Brambl, R. In press (1994).
Gavrias, V., Cubero, B., Cazelle, B., Sophianopoulou, V., and Scazzocchio, C. The proline utilisation gene cluster of *Aspergillus nidulans*. In: The Genus *Aspergillus*: From Taxonomy to Industrial Application, edited by K.A. Powell, J.F. Peberdy, and A. Renwick. FEMS Symposium No. 69 pp. 225-232 (1994), Plenum Press.
Hondmann, D.H.A. and Visser, J. (1994): Carbon metabolism in Aspergilli In: Martinelli, J.D. and Kinghorn, J.R. (eds) *Aspergillus*: 50 years on, Elsevier, Amsterdam, pp. 61-140.
Kulmburg, P., Mathieu, M., Dowzer, C., Kelly, J., and Felenbok, B. Specific binding sites in the *alc*R and *alc*A promoters of the ethanol regulon for the CREA repressor mediating carbon catabolite repression in *Aspergillus nidulans*. Mol. Microbiol., **7**, 847-857 (1993).
Peñalva, M.A., Espeso, E., Pérez-Esteban, B., Orejas, M., Fernández-Cañón, J.M. and Martinez-Blanco, H. (1993) Expression of fungal genes involved in penicillin biosynthesis. World of Biotechnol. **9**, 461-467.
Pérez-Esteban, B., Orejas, M., Gómez-Pardo, E. and Peñalva, M.A. (1993) Molecular characterisation of a fungal secondary metabolism promoter: transcription of the *Aspergillus nidulans* isopenicillin N synthesase gene is modulated by upstream negative elements. Mol. Microbiol. **9**: 881-895.
Scazzocchio, C. (1994). The proline utilisation pathway, history and beyond. In: *Aspergillus*: 50 years on. edited by Martinelli, S. and Kinghorn, J.R., Elsevier Scientific, Amsterdam, pp. 259-278.
Sequeval, D., and Felenbok, B. Relationships between zinc content and DNA-binding activity of the DNA-binding motif of the transcription factor ALCR in *Aspergillus nidulans*. Mol. Gen. Genet., **241**, 33-39 (1994).
Vries de, R.P., Flipphi, M.J.A., Witteveen, C.F.B. and Visser, J. (1994)Characterization of an *Aspergillus nidulans* L-arabitol dehydrogenase mutant. FEMS Microbiol. Lett. **123**, 83-90.

Stability of genetic information in *Bacillus* (BIOT CT-910268)

COORDINATOR:
1. S.D. EHRLICH, INRA GM, Jouy-en-Josas, F
PARTICIPANTS:
2. K.M. DEVINE, TC Dublin, Dublin, IRL
3. S. BRON, Univ Groningen, Haren, NL
4. M. SALAS, Centro Biol. Mol. CSIC, Madrid, E
5. L. BOE, Technic Univ., Lyngby, DK
6. B. DIDERICHSEN, NOVO Nordisk, Bagsværd, DK

BACKGROUND INFORMATION

Various microorganisms are used by industry for a panoply of purposes, such as production of small and large molecules, alimentary and non-alimentary fermentations, bio-remediation, pest control etc. A great hindrance for such uses are variations, and in particular loss, of interesting properties of microorganisms. Such variations are frequently encountered with microorganisms that harbour foreign genetic information, introduced to confer novel and particularly useful properties upon a microorganism. This points to a great plasticity of the microbial genetic information and raises questions about the processes which underlie this plasticity and about the ways they may be counteracted. The research undertaken under the present contract aimed to address these questions within the context of a model organism, *Bacillus subtilis*, which is a Gram positive bacterium representative of many industrial microorganisms.

OBJECTIVES AND PRIMARY APPROACHES

Foreign genetic information is most often introduced into microorganisms on extrachromosomal elements. The research undertaken was therefore focused on parameters that affect plasmid stability. However, processes that affect the integrity of DNA in general were also intensely studied. The studies ranged from theoretical considerations on the appearance of mutations and of plasmidless cells in the bacterial population, to development of systems preventing growth of cells that have lost the plasmids. They included analyses of DNA replication, carried out *in vivo* with plasmids that use rolling circle or theta replication and *in vitro* with purified proteins. Effects of DNA structure, arrest of DNA replication and DNA transcription on DNA integrity were also studied. Short term objectives were to develop systems and approaches allowing to identify and determine the importance of various parameters affecting stability of genetic information in bacteria. Long term objectives were to develop systems and approaches to control the stability of genetic information.

RESULTS AND DISCUSSION

A. Appearance of mutations in bacterial population

Studies at the Polytechnical University indicate that the current mathematical models describing bacterial populations subject to mutation are inadequate. Classical statistical analysis and fluctuation analysis of point mutations to nalidixic acid resistance and reversion of the amber mutation in the *lacZ* gene show that calculation of the mutation rates by such models is in error, or at least doubtful. Similar study of plasmid loss gave no fit between the experimental observations and the theory. Apparently, a sub-population of plasmid-containing cells gives rise to

progeny that produces plasmid-free cells at a high and unpredictable rate. Two reasons for the lack of fit were identified:

(i) host mutations with impaired plasmid segregation and
(ii) plasmid multimerization resulting in high loss rates.

The latter problem can be dealt with by introducing a multimer resolution system on the plasmid. Studies at INRA have shown that interactions between a parental plasmid and a re-arranged progeny plasmid greatly affect the appearance of cells that carry the recombinant plasmid. Two classes of interactions were identified, one affecting the establishment of the progeny plasmid the other affecting the expression of a gene on the progeny plasmid.

In both cases, decrease in the copy number of the parental plasmid favoured the appearance of cells carrying the recombinant plasmid. This indicates that the change in bacterial population from the one harbouring the desired plasmid to the one harbouring its undesired variant might be prevented by maintaining the desired plasmid at a sufficiently high copy number.

B. Characterisation of DNA replication

1. Rolling circle plasmids

A comparative analysis of the pC194 family of rolling circle plasmids was carried out at Trinity College. Plasmids of this family have been isolated from widely divergent species of bacteria, including Gram positives and Proteobacteria, Spirochaetes and Cyanobacteria. The replication (Rep) proteins of plasmids from Gram positives fall in five groups, two of which contain plasmids from *Staphylococcus aureus* . This indicates that there has been at least one transfer of plasmids between divergent Gram positive species. Double stranded origins (dso) of plasmids from Gram positive bacteria are all similar, but differ from origins of plasmids from Spirochaeta and Cyanobacteria. This indicates that the similar Rep proteins can act on a variety of substrates. Single- stranded origins (sso) are not conserved and could be critical in determination of the plasmid host range. A novel family of the rolling circle plasmids was identified at the Polytechnical University.

Mutation analysis of the replication initiator protein of plasmid pC194, RepA, was undertaken at INRA, targeting tyrosine and acidic amino acids which are well conserved among numerous related plasmids. One tyrosine and two glutamate residues were found to have a catalytic role. Tyrosine residues were previously found in active sites of different rolling circle replicons and topoisomerases, but not in association with acidic residues, which are a hallmark of DNA hydrolysing enzymes, such as exo- and endonucleases. This indicates that the active site of RepA contains two different catalytic centres, corresponding to a tyrosine and a glutamate. The former might be involved in the formation of the covalent DNA-protein intermediate at the initiation step of rolling circle replication, the latter may catalyse the release of the protein from the intermediate at the termination step. The release prevents re-initiation of plasmid replication, which is imperative for regulation of plasmid copy number by controlling the synthesis of the Rep protein. In contrast, re-initiation of another rolling circle replicon from the pC194 family, phage ΦX174, readily takes place. The phage Rep protein contains two tyrosines in the active site, which bind alternatively to DNA. Replacement of the glutamate by tyrosine in the active site of the plasmid Rep protein allowed re-initiation, which confirms, on the one hand the role attributed to the glutamate and on the other, the close relation of the phage and plasmid Rep proteins.

Detailed studies of the sso of the plasmid *Bacillus* pBAA1, carried out at the Trinity College, identified the elements important for the origin activity. They showed that RNA polymerase is involved in initiation of DNA synthesis, which is typical for rolling circle sso. Alternative pathways of initiation were searched at Groningen University, using two model plasmids, pMV158 from *Streptococcus* and pWV01 from *Lactococcus*. Both plasmids have a broad host range and replicate in *B. subtilis* and *E. coli*. The sso of pWV01 was delineated to a 250 bp segment, containing two inverted repeats, which are required for optimal conversion of single-stranded (ss) DNA to a double-stranded (ds) form and segregational stability. The conversion was carried out by RNA polymerase if both repeats were present, but could also occur, albeit at a lower efficiency, without RNA polymerase. This conversion required only one of the repeats and might be dependent on either the host primase or the primosome complex. Conversion of a derivative of pMV158, designated pLW3Em, was found to be dependent on the host RecA protein and a small complementary RNA. This indicates that initiation of DNA synthesis may require annealing of the RNA on DNA template, promoted by the activity of the RecA protein.

2. Theta replicating plasmids

Replication of a broad host range plasmid pAMβ1, which is a representative of a large plasmid family, was studied in some detail *in vivo* and *in vitro* at INRA. This plasmid is particularly interesting since it is a more stable DNA cloning vector in *B. subtilis* than the rolling circle plasmids. The following model of pAMB1 was elaborated.

(i) An RNA transcript is cleaved at the origin of DNA replication. The origin is localised 27 bp downstream from the end of the plasmid replication protein (Rep) and it is thought that the transcript cleaved is that of the Rep protein.
(ii) The cleaved transcript is used as primer by DNA polymerase I (Pol I) for initiation of DNA synthesis.
(iii) The polymerase copies the template strand and displaces the complementary strand until it is arrested by the plasmid resolvase bound to its site. The arrest takes place about 225 bp from the origin. A primosome assembly site, present on the complementary strand, is thus rendered single-stranded.
(iv) The primosome is assembled and its helicase activity displaces the resolvase from its site. Its primase activity generates primers for initiation of the lagging strand synthesis.
(v) DNA polymerase III (Pol III) loads at the end of the DNA strand liberated by the removal of the resolvase and completes the DNA synthesis.

Regulation of the pAMβ1 copy number was also studied at INRA. Two different types of regulation were uncovered, both acting at the level of transcription. One corresponds to a classical repressor operator system, the other to a counter-transcript-driven transcriptional attenuation system. Regulation occurs by controlling the synthesis of the Rep transcript, which might either limit the amount of the Rep protein or the primer for initiation of DNA replication.

Replication regions of several large plasmids were cloned at the Trinity college, but displayed a high level of structural instability, which precluded their further characterisation. The minimal replicon of the large (65 kb) conjugative *B. subtilis* plasmid pLS20 was analysed at the Groningen University. The minimal replicon is rich in dyad symmetry and does not contain any open reading frames >70 codons. It does not require DNA Pol I or RecA for replication. In the vicinity of the minimal replicon a sequence resembling the chromosomal replication terminator is present.

The sequence binds the host replication terminator protein (Rtp) and arrests replication in both orientations. These results indicate that the replicon belongs to a novel plasmid family. The minimal replicon is lost from cells at a high frequency and may be an attractive vector for the selection of stability functions.

3. Bacteriophage Φ29 All of the studies were carried out at CSCI.

Extensive analysis of Φ29 DNA polymerase by mutagenesis and biochemical characterisation of the purified mutant proteins led to identification of the role of different conserved amino acid motifs. The region 1 (motif DX_2SLYP) is involved in metal and dNTP binding, the region 2a (Kx_3NSxYG) is involved in template-primer binding and dNTP selection, the region 2b (Tx_2GR) is involved in the stabilisation of the primer terminus and region 4 (motif KxY) is involved in primer utilisation. Exonuclease activity of the enzyme was analysed in a similar way. Critical functions of four carboxylic residues and a secondary role of a tyrosine were identified by the effect of mutations on the catalytic rate. Three of the carboxylic residues were shown to be important for strand displacement, suggesting that the exonuclease and the strand displacement domains overlap. Studies of the phage p6 protein have uncovered a new motif for binding to DNA through the minor groove. Analysis of p6-DNA complexes *in vitro* indicates that they may have a role in the organisation of the Φ29 genome into a nucleoid-type nucleoprotein structure.

A mutational analysis of the Φ29 DNA right end indicates that:

(i) there are no strict sequence requirements for protein-primed initiation on single-stranded DNA;
(ii) initiation of replication occurs opposite the second nucleotide at the 3' end of the template;
(iii) a terminal repetition of a least two nucleotides is required to efficiently elongate the initiation complex; and
(iv) all the nucleotides of the template, including the 3' terminal one, are replicated. These results can be accounted for by a sliding-back model, in which a special transition step occurs between initiation and elongation of DNA replication.

Using four Φ29 proteins involved in replication (terminal protein, DNA polymerase, p5 protein (SSB) and p6 protein) it was possible to amplify the 19 kb phage DNA by 3 orders of magnitude within 1 hr at 30°C. The synthetic DNA was as biologically active as the DNA prepared from virions, which shows a high fidelity of *in vitro* amplification.

C. Factors affecting integrity of DNA

1. Recombination between short direct repeats

About a half of reported deletions occur by recombination between short directly repeated sequences. The effect of distance between the repeats on deletion frequency was examined at INRA, in a *B. subtilis* plasmid and the chromosome. The frequency decreased exponentially >1000-fold as the distance increased from 30 to 2300 bp. The decrease occurred in two phases, which might be determined by the flexibility of the DNA duplex.

Mechanism of recombination between the direct repeats was investigated at INRA, both *in vivo* and *in vitro*. The first type of studies was undertaken in *E. coli*, using the nearly precise excision of the transposon related to *Tn*10 as a model system. The excision took place from a plasmid that could replicate in the theta mode or in the rolling circle mode. It was stimulated up to 1000-fold by induction of rolling

circle replication and did not involve transfer of DNA from the parental to the progeny molecules. This indicates that it occurred by a copy-choice mechanism of DNA recombination. The studies *in vitro*, carried out with the *E. coli* DNA polymerase III, Klenow fragment of the DNA polymerase I and T4 DNA polymerase, using a single-stranded DNA template, confirmed that recombination between short direct repeats placed in the vicinity of long inverted repeats can occur by a copy choice mechanism. This mechanism was not observed, however, with the DNA polymerase of phage Φ29 and T7, indicating that different enzymes may have different requirements for slippage on a single-stranded DNA template.

2. Recombination between non-homologous sequences

About one half of reported deletions occur between sequences of no homology and are thought to be due to errors of DNA cut and join enzymes. To examine the factors which might affect this process a model system was used at INRA, consisting of hybrids between a theta-replicating plasmid and a rolling circle replicating bacteriophage, which undergo deletions in *E. coli* with a very high frequency. One deletion point is always at the nick site in the rolling circle replication origin, introduced by the replication protein, and the other end point was found to be almost exclusively within the region which is rendered positively supercoiled by convergent transcription. When the transcription is repressed, the other deletion end-point is often found at the operator site, provided that the direction of the plasmid replication progressed towards the operator. This indicates that DNA topology and the arrest of DNA replication by the repressor-operator interaction greatly affect deletion formation.

To study the involvement of topoisomerases in deletion formation the analysis of the *B. subtilis* gene encoding topoisomerase I (*top*I) was undertaken at Groningen University. The gene was cloned by PCR techniques and its sequence was determined. It contains three zinc fingers and can complement *E. coli top*A mutants. Deletion of the last finger had no effect on complementation, but the deletion of all three fingers led to a loss of activity. The gene could not be inactivated in *B. subtilis*, which indicates that it has an essential role. It is expressed constitutively, albeit at a low level, and does not respond to high osmotic pressure, as judged by fusions with the *lacZ* reporter gene. The effect of the gene over-expression on plasmid stability is being investigated.

3. Stability of the chromosomal genes

Activation of a RCR origin in the *B. subtilis* chromosome stimulates deletion and amplification of repeated sequences in its vicinity. To study the effect of activation of a theta origin pAMβ1 replicon was inserted in the chromosome. The origin was not functional at the high temperature (51°) and was activated at low temperature (37°), as shown by electrophoretic and electron microscope analysis. Both recombination and amplification were only marginally affected by origin activation. This indicates that the type of replication a genome undergoes influences highly its stability.

D. Post-segregational plasmid stability system

The studies were carried jointly at the Trinity College and NOVO-Nordisk. The system is based on the expression of the lysis genes of PBSX, located on the *B. subtilis* chromosome while the repressor of lysis gene expression resides on the plasmid to be stabilised. Loss of the plasmid will thus result in expression of the lysis functions and cell death.

The lysis genes of PBSX were located to a region near the end of the late operon, on a 3.6 kb fragment of DNA. This fragment was sequenced and found to contain four open reading frames; ORF1 is 279 amino acids in length and shows homology to an incomplete open reading frame found upstream of an autolytic amidase (*cwl*A) from *B. subtilis*; ORF2 is 89 amino acids in length, contains a hydrophilic amino terminal region and a hydrophobic carboxy terminal region and is of unknown function. ORF3 is 87 amino acids in length and is structurally similar to the holin family of proteins. It contains two hydrophobic regions, each of which can form a trans-membrane alpha helix, separated by a short hydrophilic region containing the helix disrupting amino acids glycine and proline. ORF4 encodes a protein of 298 amino acids in length which is homologous to *cwl* gene which encodes the autolytic amidase in *B. subtilis*. Mutation analysis indicated that neither the amidase alone, nor the amidase and the holin protein are sufficient for cell lysis and point to ORFs 1, 2 and 3 as playing a critical role in the lytic process.

A *B. subtilis* strain was constructed in which the four lysis genes are juxtaposed to the transcription initiation point of the late operon and which also contains a temperature sensitive allele of the phage repressor *xre*-1, so that lysis can be induced by a temperature upshift to 48°C. The strain was made *rec*A to prevent recombination between the plasmid which carries the wild-type repressor and the chromosome which carries a temperature sensitive allele of the repressor. The plasmids pHV1432 and pHV1432/xre, which carries the wild-type PBSX repressor were transformed into this strain. The stability of these plasmids will be examined in cells grown at 37°C and at 48°C. At 37°C there will be no selection for the retention of the plasmid in the cell population since the chromosomal copy of the repressor is functional. In contrast, at 48°C the chromosomal copy of the repressor is inactive and cell survival is expected to depend on maintaining the plasmid. A detailed kinetic analysis of plasmid loss at either temperature is being carried out, and the system is being modified to become more mobile.

MAJOR SCIENTIFIC BREAKTHROUGHS

A discrepancy between the theory and observation of mutant appearance in bacterial populations was discovered. A plasmid replicating by a novel mechanism was found. A chimeric structure and function of a rolling circle plasmid replication protein, using topoisomerase and nuclease activities was revealed. A novel mechanism of initiation and elongation of DNA replicaiton was identified. A novel mechanism of initiation of phage DNA replication, designated sliding-back, was discovered. Rapid and efficient system for amplification of large DNA segments was established. A mechanism of recombination between short direct repeats was identified. *B. subtilis* phage lysis system was unravelled and used to stabilise plasmids by post-segregational killing.

MAJOR COOPERATIVE LINKS

Two meetings of all participants have been organized, in Jouy en Josas in 1992 and Madrid in 1993; a third took place in Groningen in September 1994. Moreover, several meetings between two or more of the participating laboratories were organized, as well as visits of individual scientists. Collaborative efforts are being carried out by INRA and the University of Groningen. Contacts are maintained between different laboratories by telephone, mail and the exchange of materials. Close contacts between the T project on lactic bacteria and the present program are entertained, via the University of Groningen and the INRA laboratories.

PUBLICATIONS

Joint publications

S. de Jong, P. Noirot, S.D. Ehrlich, G. Venema, S. Bron (1994) Cloning and characterization of the *Bacillus subtilis* topoisomerase I gene. (in preparation).
S. Krogh, S. Jørgensen, B. Diderichsen, K.M. Devine (1994) Analysis of the lysis genes contained on the phage PBSX. (in preparation).

Individual publications

E. d'Alençon, M. Petranovic, B. Michel, P. Noirot, A. Aucouturier, M. Uzest, S.D. Ehrlich (1994) Copy choice illegitimate DNA recombination revisited. EMBO J. **13**, 2725-2734.
L. Andrup, J. Damgaard, K. Wassermann, L. Boe, M.M. Madsen, F.G. Hansen (1994) Complete nucleotide sequence of the *Bacillus thuringiensis* subsp. *israeliensis* plasmid pTX14-3 and ist correction with biological properties. Plasmid **31**, 72-88.
V. Bidnenko, A. Gruss, S.D. Ehrlich (1993) Mutation in the plasmid pUB110 Rep protein affects its ability to terminate rolling circle replication. J. Bacteriol. **175**, 5611-5616.
H. Bierne, S.D. Ehrlich, B. Michel (1994) Flanking sequences affect replication arrest at the *Escherichia coli* terminator *TerB in vivo*. J. Bacteriol., **176**: 4165-4167.
H. Bierne, S.D. Ehrlich, B. Michel (1994) Replication terminator facilitates detection of plasmids formed by recombination between short repeats in *Escherichia coli*. Plasmid, (in press).
I. Biswas, V. Vagner, S.D. Ehrlich (1992) Efficiency of homologous intermolecular recombination at different locations of the *Bacillus subtilis* chromosome. J. Bacteriol. **174**, 5593-5596.
L. Blanco, M. Salas (1994) Mutational analysis of bacteriophage Φ29 DNA polymerase. Methods Enzymol., (in press).
L. Blanco, J.M. Lazaro, M. de Vega, A. Bonnin, M. Salas (1994). *In vitro* amplification of infective phage Φ29 DNA. Proc. Natl. Acad. Sci. USA, in press.
M.A. Blasco, J.M. Lazaro, A. Bernad, L. Blanco, M. Salas (1992) Φ29 DNA polymerase active site: mutants in conserved residues Tyr^{254} and Tyr^{390} are affected in dNTP binding. J. Biol. Chem. **267**, 19427-19434.
M.A. Blasco, J.A. Esteban, J. Mendez, L. Blanco, M. Salas (1992) Structural and functional studies on Φ29 DNA polymerase. Chromosoma **102**, 532-538.
M.A. Blasco, J.M. Lazaro, L. Blanco, M. Salas (1994) Φ29 DNA polymerase active site. The conserved amino acid motif KX_3NS_xIG is involved in metal and dNTP binding. J. Biol. Chem., **268**, 16763-16770.
M.A. Blasco, J.M. Lazaro, L. Blanco, M. Salas (1993) Φ29 DNA polymerase active site. Residue Asp^{249} of conserved amino acid motif Dx_2SLYP is critical for synthetic activities. J. Biol. Chem. **268**, 24106-24113.
L. Boe (1992) Translational errors as the cause of mutations in *Escherichia coli*. Mol. Gen. Genet. **231**, 469-471.
L. Boe, T. Tolker-Nielsen, K.M. Eegholm, H. Spliid, A. Vrang (1994) Fluctuation analysis of mutation to nalidixic acid resistance in *Escherichia coli*. J. Bacteriol. **176** (10). (in press).
A. Bravo, J.M. Hermoso, M. Salas (1994) A genetic approach to identify functionnal amino acids in protein p6 of *Bacillus subtilis* phage Φ29. Mol. Gen. Genet. (in press).
A. Bravo, J.M. Hermoso, M. Salas (1994) *In vivo* functionnal relationships among terminal proteins of *Bacillus subtilis* Φ29-related phages. Gene. **148**, 197-112.
C. Bruand, E. Le Chatelier, S.D. Ehrlich, L. Janniere (1993) A novel class of theta-replicating plasmids — the pamβ1 family from Gram-positive bacteria. Proc. Natl. Acad. Sci. USA **90**, 11668-11672.
C. Bruand, S.D. Ehrlich, and L. Janniere (1994) Lagging-strand initiation site of the *Enterococcus faecalis* plasmid pAMβ1. (submitted).
F. Chédin, E. Dervyn, R. Dervyn, S.D. Ehrlich, P. Noirot (1994) Frequency of deletion formation decreases exponentially with distance between short direct repeats. Molec. Microbiol. **12**, 561-569.
C. Cordes, B. Twiest, B. Kazemier, G. Venema, J.M. van Dijl, S. Bron (1994) Possible role of DNA supercoiling in structural plasmid instability caused by divergent transcription and secretion of a plasmid- specified protein in *Bacillus subtilis*. (in preparation).
P. Dabert, S. Sourice, S.D. Ehrlich, A. Gruss (1994) Analyses of ψ and *recBCD* alleles by rolling circle plasmids in *Escherichia coli*. (submitted).

S.D. Ehrlich, H. Bierne, E. d'Alencon, D. Vilette, M. Petranovic, P. Noirot and B. Michel (1994) Mechanisms of illegitimate recombination. Gene **135**, 161-166.
J.A. Esteban, M. Salas, L. Blanco (1993) Fidelity of Φ29 DNA polymerase. Comparison between protein- primed initiation and DNA polymerization. J. Biol. Chem. **268**, 2719-2726.
J.A. Esteban, M.S. Soengas, M. Salas, L. Blanco. 3'->5' exonuclease active site of Φ29 DNA polymerase. Evidence favoring a metal-ion-assisted reaction mechanism. (1994). J. Biol. Chem., (in press).
R. Freire, M. Salas, J.M. Hermoso. A new protein motif for binding to DNA through the minor groove. EMBO J. **13**, 4353-4360.
C. Gutierrez, R. Freire, M. Salas, J.M. Hermoso (1994) Assembly of phage Φ29 genome with viral protein p6 into a compact complex. EMBO J. **13**, 269-276.
J.M. Hermoso, R. Freire, A. Bravo, C. Gutierrez, M. Serrano, M. Salas (1994) DNA structure in the nucleoprotein complex that activates replication of phage Φ29. Biophys. Chem., **50**, 183-189.
S. de Jong, G. Venema, S. Bron (1994) Structural analysis of the *Bacillus subtilis* topoisomerase I gene product. (in preparation).
L. Jannière, A. Gruss, S.D. Ehrlich (1993) Plasmids from Gram positive bacteria. In 'Bacillus and other Gram-positive bacteria: Biochemistry, physiology and molecular genetics' (J.A. Hoch, A.L. Sonensheim & R. Losick eds.), American Society for Microbiology, Washington DC, pp. 625-644.
J.M. Lazaro, L. Blanco, M. Salas (1994) Purification of bacteriophage Φ29 DNA polymerase. Meth. Enzymol., (in press).
E. Le Chatelier, S.D. Ehrlich, L. Janniere (1993) Biochemical and Genetic analysis of the unidirectional theta replication of the *S. agalactiae* plasmid pIP501. Plasmid **29**, 50-56.
E. Le Chatelier, S.D. Ehrlich, L. Janniere (1994) The pAMB1 CopF repressor regulates plasmid copy number by controlling transcription of the *repE* gene. Molec. Microbiol., (in press).
E. Le Chatelier, S.D. Ehrlich, L. Janniere (1994) Countertranscript-driven transcriptional attenation in the pAMβ1 *repE* gene. (submitted).
S.M. Madsen, L. Andrup, L. Boe (1993) Fine mapping and DNA sequence of replication functions of *Bacillus thuringiensis* plasmid pTX14-3. Plasmid, **30**, 119-130.
W.J.J. Meijer, D. van der Lelie, G. Venema, S. Bron (1994) Effects of the generation of single-stranded DNA and the presence of a mobilization function on the maintenance of plasmid pMV158 and derivatives in *Lactococcus lactis*. (submitted).
W.J.J. Meijer, D. van der Lelie, G. Venema, S. Bron (1994) Effects of the generation of single-stranded DNA on the maintenance of plasmid pMV158 and derivatives in different *B. subtilis* strains. (submitted).
W.J.J. Meijer, G. Venema, S. Bron (1994) Selection and characterization of single-strand origins from endogenous *Bacillus subtilis* plasmids. (in preparation).
W.J.J. Meijer, S. van Tongeren, A. de Boer, M. Smith, G. Wake, G. Venema, S. Bron (1994) Cloning and characterization of the minimal replicon of the large *Bacillus subtilis* plasmid pLS20 which contains a replication terminator. (in preparation).
R. Meima, B.J. Haijema, H. Dijkstra, G. Venema, S. Bron (1994) Structural plasmid instability in *Bacillus subtilis*: recombinational repair as a source of plasmidic rearangements. (in preparation).
R. Meima, G. Venema, S. Bron (1994) A positive selection vector for the analysis of structural plasmid instability in *Bacillus subtilis*. (in preparation).
J. Mendez, L. Blanco, J.A. Esteban, A. Bernad, M. Salas (1992) Initiation of Φ29 DNA replication occurs at the second 3' nucleotide of the linear template: A sliding-back mechanism for protein-primed DNA replication. Proc. Natl. Acad. Sci. USA **89**, 9579-9583.
S. Molin, L. Boe, L.B. Jensen, C.S. Kristensen, M. Givskov, J.L. Ramos, A.K. Bej (1993) Suicidal genetic elements and their use in biological containment of bacteria. Ann. Rev. Microbiol. **47**, 139-166.
M.F. Noirot-Gros, V. Bidnenko, S.D. Ehrlich (1994) Active site of the replication protein of the rolling circle plasmid pC194. EMBO J., **13** 4412-4420.
C. Pujol, S.D. Ehrlich, L. Janniere (1994) The promiscous plasmids pIP501 and pAMβ1 from Gram-positive bacteria encode cross-reacting resolution functions. Plasmid **31**, 100-105.
C. Pujol, S.D. Ehrlich, L. Janniere (1994) Inhibition of a naturally integrated RCR replicon in derivatives of *the Streptococcus agalactiae* plasmid pIP501. (submitted).

M.S. Soengas, J.A. Esteban, J.M. Lazaro, A. Bernad, M.A. Blasco,M. Salas and L. Blanco (1992) Site-directed mutagenesis at the Exo III motif of Φ29 DNA polymerase; overlapping structural domains for the 3'-> 5'exonuclease and strand-displacement activities. EMBO J. **11**, 4227-4237.
J.F.M.L. Seegers, A.C. Zhao, S.A. Khan, L.E. Pearce, W.J.J. Meijer, G. Venema and S. Bron (1994) Lagging strand synthesis via the single-strand origin from the lactococcal plasmid pWV01 can occur in a rifampicin insensitive manner. (submitted).
L. Seery and K.M. Devine (1993) Analysis of the features contributing to the activity of the origin of the single-stranded origin of the *Bacillus* plasmid pBAA1. J. Bacteriol. **175**, 1988-1994.
L. Seery, N.C. Nolan, P.M. Sharp, K. Devine (1993) Comparative analysis of the pC194 group of rolling circle plasmids. Plasmid **30**, 185-196.
M. Serrano, C. Gutierez, M. Salas, J.M. Hermoso (1993) The superhelical path of the DNA in the nucleo- protein complex that activates the initiation of phage Φ29 DNA replication. J. Mol. Biol., **230**, 248-259.
M. Serrano, M. Salas, J.M. Hermoso (1993) Multimeric complexes formed by DNA-binding proteins of low sequence-specificity. TIBS **18**, 202-206.
M. Serrano, C. Gutierrez, R. Freire, A. Bravo, M. Salas, J.M. Hermoso (1994) Phage Φ29 protein p6: a viral histone-like protein. Biochimie, (in press).
M.S. Soengas, J.A. Esteban, M. Salas and C. Gutierrez (1994) Complex formation between phage Φ29 single-stranded DNA binding protein and DNA. J. Mol. Biol., **239**, 213-226.
D. Vilette, M. Uzest, S.D. Ehrlich, B. Michel (1992) DNA transcription and repressor binding affect deletion formation in *Escherichia coli* plasmids. EMBO J. **11**, 3639-3644.
T. Tolker-Nielsen and L. Boe (1994). A statistical analysis of the formatiuon of plasmid-free cells in populations of *Escherichia coli*. J. Bacteriol. **176**, 4306-4310.
P. Ravn, M. Givakov, K.-M. Eegholm, S.M. Madsen and L. Boe (1994). Observations on the formation of deletions on monomeric and dimeric plasmids in *Escherichia coli*. Molec. Microbiol. **14**, (in press).

Physiology and molecular genetics of amino acid production and secretion by corynebacteria: flux of intermediates and feedback control mechanisms (BIOT CT-910264)

COORDINATOR:
J.F. MARTÍN, Universidad de León, León, E
PARTICIPANTS:
H. SAHM, Forschungszentrum Jülich GmbH, Jülich, D
K. DUNICAN, University College, Galway, IRL
A. PÜHLER, Universität Bielefeld, Bielefeld, D
C. BLANCO, Université de Rennes I, Rennes, F

BACKGROUND INFORMATION

Industrial production of amino acid (glutamic acid, lysine, threonine, tryptophan, phenylalanine, etc) is one of the main areas of Biotechnology, both in Europe and in the rest of the world, particularly in Japan and in USA. Significant improvements in amino acid yields can be obtained by
i) increasing the uptake of nutrients by the producer microorganisms,
ii) directing the flow of intermediates of primary metabolism,
iii) overexpressing the genes that lead to the biosynthesis of the amino acids and
iv) increasing the efficiency of the export systems.

OBJECTIVES AND PRIMARY APPROACHES

I. Development of basic molecular tools
1. Conjugal gene transfer (Pühler)
2. Gene disruption and gene replacement (Pühler)
3. Transposon mutagenesis (Pühler)
4. Construction of an encyclopedia of *Brevibacterium lactofermentum* and *Corynebacterium glutamicum* (Martín)
5. Mapping of the *Corynebacterium glutamicum* chromosome (Pühler)
6. Characterization of a region of pBL1 of *Brevibacterium lactofermentum* involved in plasmid replication (Martín)
7. RNA polymerase sigma factors in corynebacteria (Martín)
8. Phage-derived integrative vectors (Blanco)
9. Construction of broad host range integrating plasmids for the genetic manipulation of selected bacterial strains (Dunican)

II. Industrial applications: cloning and analysis of amino acid biosynthesis genes
10. Transcriptional analysis of the *hom-thrB* cluster (Martín, Sahm)
11. Molecular analysis of aromatic amino acid biosynthetic genes in *Corynebacterium glutamicum* (Dunican)

III. Biochemical studies
12. Characterization of central pathways and TCA cycle key enzymes in *C. glutamicum* (Sahm)
13. Characterization of the glyoxylate cycle enzymes in *C. glutamicum* (Sahm)

RESULTS AND DISCUSSION

I. Development of basic molecular tools

1. *Conjugal gene transfer* (Pühler)

A highly efficient conjugal transfer system between *E. coli* and *C. glutamicum* has been developed. We were able to complement a restriction-deficient *C. glutamicum* mutant with a DNA fragment from the wild type chromosome. The DNA fragment was shown also to confer resistance against phage CL31 when grown on a heterologous host. The DNA fragment was sequenced and four open reading frames were identified. Subcloning analysis identified ORF1 as responsible for complementation of the restriction-deficiency. The ORF was disrupted in the *C. glutamicum* wildtype strain ATCC 13032 and yielded a restriction-minus phenotype. It was shown that ORF1 encodes a restriction enzyme that specifically cuts heterologous DNA. The enzyme shows ATP-dependence and heat-sensitivity. It is concluded that the enzyme *CglI* is responsible for the high conjugal competence after heat treatment of *C. glutamicum*.

2. *Gene disruption and gene replacement* (Pühler)

Two new vector systems for the conjugal transfer of non-replicating vectors to corynebacteria were developed. The first system is effective in the selection for double crossover events that lead to marker exchange between a plasmid and the chromosome. The double crossover was selected for by including the *Bacillus subtilis sac*B gene in the vector. *Sac*B expression is lethal in *C. glutamicum* when the cells were grown on sucrose.

The second vector includes the promoterless *E. coli* lacZ gene coding for β-galactosidase as a reporter gene. Transcriptional fusions between genes and the reporter plasmid were induced by the gene disruption method. This technique has the advantage that the transcriptional activity was determined in the single copy state within the chromosome.

3. *Transposon mutagenesis* (Pühler)

Several transposon mutants of *C. glutamicum* generated by the transposon TnCx1 were investigated by cloning and sequencing of the insertion sites and by mapping of the insertions against large chromosomal fragments separated by pulsed-field gel electrophoresis.

4. *Construction of an encyclopedia of* **Brevibacterium lactofermentum** *and* **Corynebacterium glutamicum** (Martín)

The chromosomal DNA was resolved in a discrete number of fragments using restriction endonucleases that recognize AT-rich hexanucleotide or octanucleotide sequences. Using SwaI (recognition sequence ATTTAAAT) the genome of *B. lactofermentum* 13869 (total size 3031 kb) was resolved in 19 fragments of 18 to 550 kb; the genome of *B. lactofermentum* DSM 20412 in 21 fragments and the genome of *C. glutamicum* ATCC 13032 (total size 2958 kb) in 26 fragments. Similarly, using *Pac*I (recognition sequence TTAATTAA) the genome of ATCC 13869 was resolved into 26 fragments with a size range of 5 to 390 kb amounting to 3026 kb, *B. lactofermentum* DSM 20412 into 23 fragments with a total size of 3044 kb and *C. glutamicum* ATCC 13032 into 27 fragments with total genome size of 2987 kb. A large number of markers have been mapped in the DNA fragments resolved by PFGE and in cosmids by hybridizations with probes internal to those genes.

5. ***Mapping of the* Corynebacterium glutamicum *chromosome*** (Pühler)

The map of the *C. glutamicum* ATCC 13032 chromosome was worked out further. The genome size was determined to be 3080 +/- 15 kilobases by cutting with the meganucleases SwaI, PacI and PmeI. From a cosmid library, clones were selected that span SwaI sites. By hybridization with these cosmids we were able to link 15 fragments to two large contigs covering 80% of the genome. The rest of the fragments will be linked by cross-hybridization to PacI- and PmeI- fragments.

The physical mapping of the *C. glutamicum* ATCC 13032 chromosome is near to completion (95% of the fragments are linked) and 576 cosmids are sorted by the use of hybridization techniques. The fingerprint technique was worked out and initially used to link cosmids of the largest SwaI fragment of the chromosome.

6. ***Characterization of a region of plasmid pBL1 of* Brevibacterium lactofermentum *involved in replication via the rolling circle model*** (Martín)

The minimal region for autonomous replication of pBL1, a 4.5 kb cryptic plasmid of *Brevibacterium lactofermentum* ATCC 13869 that has been used to construct a variety of corynebacterial vectors, was shown to be contained on a 1.8 kb *Hind*III-*Sph*I DNA fragment. This region constains two open reading frames (ORFs) (ORF1 and ORF5) which are essential for pBL1 replication in *B. lactofermentum*. Accumulation of single-strand intermediates in some of the constructions indicates that plasmid pBL1 replicates via the rolling circle replication model; its plus strand and minus strand were identified by hybridization with two synthetic oligonucleotide probes complementary to each pBL1 strand. ORF1 seems to encode the Rep protein and showed partial homology with sequences for Rep proteins from *Streptomyces* plasmids which replicate via rolling circle replication such as pIJ101, pSB24, and pJV1.

7. ***RNA polymerase sigma factors in corynebacteria*** (Martín)

At least two different types of promoters are known in corynebacteria: *E. coli*-like promoters (CEP or type I), and corynebacteria specific promoters (CSP or type II) which are not recognized in *E. coli*. The distribution of *B. lactofermentum* sigma factor genes was studied by hybridization experiments of total DNA of *B. lactofermentum* or *C. glutamicum* with a synthetic oligonucleotide (*rpo*D probe) corresponding to the highly conserved sequences in the *E. coli* and *B. subtilis* principal sigma factors and a fragment of the cloned *Streptomyces griseus hrd*B genes (Marcos, A.T. and Martín, J.F., unpublished).

A *B. lactofermentum* cosmid library was screened with both *rpo* and *hrd*B probes, allowing the isolation of clones for three different sigma factor genes. The DNA sequences of two of these genes have been determined and the deduced protein sequences show clearly homology with the *hrd*B protein from *S. griseus* and *S. coelicolor* and with the main sigma factor genes of *E. coli, B. subtilis* and other Gram-negative and Gram-positive bacteria.

8. ***Phage derived integrative vectors*** (Blanco)

8.a. *Study of different corynephages. Characterization of corynephages isolated from soil*

Various corynephages were isolated from soil samples. Restriction analysis of phage DNA indicated that 24 phages were unique. One of them was identical to the previously characterized phage CG33. Twenty three of them are virulent phages, their hosts are essentially composed of *Arthrobacter* species.

The unique temperate phage, AAU2, made turbid plaques on *Arthrobacter aureus*. It is active only on *Arthrobacter* species. One virulent mutant was isolated that is unable to make plaques on the lysogenised strain, and thus seems affected on the repressor of lytic functions. AAU2 was characterized morphologically. It presents a polyhedral head of 50 nm wide and a non contractile tail of 150 nm, and it was classified in Ackerman group B1.

8.b. *Characterization of Corynebacterium glutamicum integrated prophages*
Twenty *C. glutamicum* strains were induced by UV and mitomycin C treatments. The presence of phages in strain lysates was searched for by infection of lawns of all strains or by DNA characterization of phage particles sedimented from strain lysates.
This screening led to the isolation of 4 distinct phages. Four *C. glutamicum* strains (ATCC 13058, ATCC 21488, ATCC 21650 and ATCC 21649) contain both Φ304L and Φ304S and two *B. divaricatum* strains (ATCC 14020 and ATCC 21792) contain both Φ15 and Φ16. They belong to the *Siphoviridae*, all having closed dimensions except Φ304S which is smaller. Their capsids contain 2 to 3 polypeptides as seen by SDS-PAGE. Their genomes have cohesive extremities.

8.c. *Characterization of integrative functions of phages ΦAAU2, Φ304S, Φ304L and Φ16*
The integrative functions of the different phages were located by subcloning DNA fragments surrounding the *att*P site, followed by restriction enzyme or exonuclease III deletions. The integrity of these functions was checked by the efficiency of integration. The minimal region necessary for the integration of the 4 phages present a similar organisation: an *att*P site and a DNA region necessary for maximal integration efficiency surrounding the integrative functions. Proteins encoded by this region were analysed in *E. coli* using a T7 RNA polymerase expression vector. The DNA sequence of the different attachement sites, and for one phage the whole *int* region was determined.
The minimal region allowing maximal integration rate was inserted into various *E. coli* plasmids. A strong influence of the *E. coli* vector on the efficiency of integration was observed: pBluescript derivatives present an integration rate 100-fold higher than p15A based derivatives. The position of int region on the plasmid has a lower influence.
Cloning vectors were constructed by inserting *int* functions on a pBluescript derivative carrying a kanamycin resistance gene. Insertion of DNA in the multiple cloning site confers a Lac- phenotype to recombinant plasmids and could be selected for in *E. coli*. In corynebacteria these constructions are non-replicative. They are able to integrate in various *Corynebacterium glutamicum* strains (ATCC 13032, 21086, 14020) insensitive to the different phages. Their efficiency of integration was very high, in some cases identical or slightly greater than the efficiency of transformation by replicative plasmids. The integrations are very stable; no loss of the resistance marker was observed over 80 generations in absence of selective pressure.

9. ***Construction of a broad host range integrating plasmids for the genetic manipulation of selected bacterial strains*** (Dunican)

Work on the construction of a broad host range integrative vector (pBC23), which is ideal for the genetic manipulation of coryneform bacteria has been completed. This plasmid which is a derivative of the conditional suicide vector pBGS8 contains a 1.5 kb fragment of the 23S rDNA operon from *C. glutamicum*, a functional *lacZ* gene and a polycloning site which facilitates cloning of genes by positive selection,

and a kanamycin gene which serves as a selectable marker in both *E. coli* and *C. glutamicum*. This plasmid has been integrated into a number of corynebacterial strains and shown to be stably mantained for many generations in the absence of antibiotic selection.

II. Industrial applications: cloning and analysis of amino acid biosynthesis genes

10. ***Transcriptional analysis of the hom-thrB cluster*** (Martín, Sahm)

Two of the genes, *hom* (encoding homoserine dehydrogenase) and *thr*B (encoding threonine kinase) of the threonine biosynthetic pathway are clustered in the chromosome of *Brevibacterium lactofermentum* in the order 5'.*hom-thr*B.3' separated by only 10 bp. The thrB gene is expressed in *E. coli*, in *Brevibacterium* and in *Corynebacterium* and complements auxotrophs of all three organisms deficient in homoserine kinase, whereas the *hom* gene did not complement two different *E. coli* auxotrophs lacking homoserine dehydrogenase. However complementation was obtained when the homoserine dehydrogenase was formed as a fusion protein in *E. coli*. Northern analysis showed that the *hom-thr*B cluster is transcribed giving two different transcripts of 2.5 and 1.1 kb. The 2.5 kb transcript corresponds to the entire cluster *hom-thr*B (i.e. they form a bicistronic operon), and the short transcript (1.1 kb) originates from the thrB gene. The promoter in front of *hom* and the *hom*-internal promoter in front of *thr*B were subcloned in promoter-probe vectors of *E. coli* and corynebacteria. The *thr*B promoter is efficiently recognized both in *E. coli* and corynebacteria whereas the promoter of the operon is functional in corynebacteria but not in *E. coli*.

11. ***Molecular analysis of aromatic amino acid biosynthetic genes in*** **Corynebacterium glutamicum** (Dunican)

We have cloned and characterized the entire tryptophan operon from the tryptophan derepressed strain *C. glutamicum* 21850. Chromosomal amplification of this construct in *C. glutamicum* 21850 has been shown to improve tryptophan productivity. In addition, all the genes from the common aromatic amino acid biosynthetic pathway have also been cloned. Three of these genes have been sequenced, namely DAHP synthase, *aro*A and *aro*D. The four remaining genes *aro*B, *aro*E, aro*L* and *aro*C have been mapped in a single DNA fragment. Subcloning and sequence analysis of these genes is currently in progress.

III. Biochemical studies

12. ***Characterization of central pathways and TCA cycle key enzymes in*** **C. glutamicum** (Sahm)

The genes encoding the *C. glutamicum* enzymes glyceraldehyde-3-phosphate dehydrogenase (*gap*), 3-phosphoglycerate kinase (*pgk*), triosephosphate isomerase (tpi), phosphoenolpyruvate carboxylase (*ppc*), citrate synthase (*gltA*), isocitrate dehydrogenase (*icd*), subunit E3 of the oxoglutarate dehydrogenase (*lpd*), isocitrate lyase (*aceA*), malate synthase (*aceA*), malate synthase (aceB) and glutamate dehydrogenase (*gdh*) have been isolated, sequenced and (over)expressed in *C. glutamicum*.
The transcriptional organization of the genes *gap, pgk, tpi, ppc, gltA, lpd* and *gdh* and the promoters of these transcriptional units have been characterized. The genes *pgk, tpi* and *ppc* form a tricistronic operon in that order with partial intercistronic termination between *tpi* and *ppc*. All other genes have a monocistronic organization.

13. *Characterization of the glyoxylate cycle enzymes in* **C. glutamicum** (Sahm)

The two key enzymes of the glyoxylate cycle, isocitrate lyase and malate synthase, were isolated from *C. glutamicum*, purified to apparent homogeneity and subjected to biochemical analysis. The native isocitrate lyase is a tetramer of identical subunits (Mr ≈ 47.200). The native malate synthase is a monomer (Mr ≈ 82.400). The *C. glutamicum* malate synthase showed high affinities to the substrates glyoxylate (30 μm) and acetyl-CoA (12 μM).

MAJOR SCIENTIFIC BREAKTHROUGHS

An impressive array of molecular genetic tools have been developed. More than 20 different genes encoding enzymes of primary metabolism and of the amino acid biosynthetic pathways have been cloned. Overexpression studies and directed removal of regulatory mechanisms have resulted in better amino acid overproducing strains.

MAJOR COOPERATIVE LINKS

Frequent meetings and scientific exchanges have been carried out between the five laboratories involved and two associated companies. Scientific meetings were carried out in León (Spain), Paris (Orsan) and Frankfurt (Degussa, 2 meetings).

Dr. Luis M. Mateos (León) spent six weeks in the lab of Dr. Pühler during August and September 1993. He was sucessful in establishing a technique to integrate mobilizable, non-replicating vectors into the corynebacterial genome in a random fashion. These vectors have no detectable homology and are probably integrated by illegitimate recombination.

Bacterial strains, phages, and plasmids were exchanged also with the groups from Galway, Ireland and Rennes, France.

Mr. Agustín Pisabarro, from the laboratory of Martín (León), stayed for two months in 1993 working with Eikmanns and Sahm at Jülich. Part of his work is enclosed in paragraph 10: Transcriptional analysis of the *hom-thrB* cluster.

Dr. Dunican made available DNA for six *Corynebacterium* genes to Dr. Pühler for use in preparing his gene map. Dr. Dunican also made available *rec* minus strains of *C. glutamicum* to Dr. Pühler and Dr. Martín.

Claire Le Marrec, from the laboratory of Blanco (Rennes), spent three months in 1994 in the lab of Dr. Pühler, working with Dr. Kalinowski on the characterization of phage integration regulation and phage-conjugative-integration vectors construction.

PUBLICATIONS

Joint publications

Mateos, L.M., Pisabarro, A, Pátek, M., Malumbres, M., Guerrero, C., Eikmanns, B.J., Sahm, H. and Martín, J.F. (1994) J. Bacteriol. (in press)

Individual publications

Barry et al. (1991) PCR Methods and Applications 1:51-56

Börmann, E.R., Eikmanns, B.J. and Sahm, H. (1992) Mol. Microbiol. 6:317-326

Börmann-El Kholy, E.R., Eikmanns, B.J., Gutmann, M. and Sahm, H. (1993) Appl. Environm. Microbiol. 59:2329-2331

Correia, A., Martín, J.F. and Castro, J.M. (1994) Microbiology 140 (in press)

Eikmanns, B.J. (1992) J. Bacteriol. 174:6076-6086

Eikmanns, B.J., Eggeling, L. and Sahm, H. (1993) Antonie van Leeuwenhoek 64:145-163

Eikmanns, B.J., Thum-Schmitz, N., Eggeling, L., Lüdtke, K. and Sahm, H. (1994) Microbiology (in press)
Fernández-González, C., Cadenas, R.F., Noirot-Gros, M.F., Martín, J.F. and Gil, J.A. (1994) J. Bacteriol. **176**:3154-3161
Fitzpatrick, R. O'Donohue, M., Joy, J., Heery, D.M. and Dunican, L.K. (1994) Appl. Microbiol. Biotechnol. (in press)
Guerrero, C., Mateos, L.M., Malumbres, M. and Martín, J.F. (1994) Gene **138**:35-41
Heery and Dunican (1990) Trends in Genetics **6**:173
Heery and Dunican (1990) Nucl. Acids Res. **18**:7138
Heery and Dunican (1993) Appl. Environm. Microbiol. **59**:791-799
Heery, D.M., Fitzpatrick, R. and Dunican, L.K. (1994) Biochem. Biophys. Res. Commun. **201**:1255-126215
Jäger, W., Schäfer, A., Pühler, A., Labes, G. and Wohlleben, W. (1992) J. Bacteriol. **174**:5462-5465
Le Marrec, C., Michotey, V., Blanco, C. and Trautwetter, A. (1994) Microbiology (in press)
Le Marrec, C., Blanco, C. and Trauwetter, S. (1994) (submitted)
Leret, V., Moreau, S., Bonnassie, S., Blanco, C. and Trautwetter, A. (1994) (submitted)
Leret, V., Le Marrec, C., Moreau, S., Bonnassie, S., Blanco, C. and Trautwetter, A. (1994) (submitted)
Malumbres, M., Gil, J.A. and Martín, J.F. (1993) Gene **134**:15-24
Malumbres, M., Mateos, L.M. and Martín, J.F. (1994) In: Food Biotechnology. Eds. Huy & Khatchatourians. VCH Publishers, Inc., Weinheim, Germany (in press)
Malumbres, M., Mateos, L.M., Lumbreras, M.A., Guerrero, C. and Martín, J.F. (1994) Appl. Environm. Microbiol. 60 (in press)
Malumbres, M., Mateos, L.M., Guerrero, C. and Martín, J.F. (1994) J. Bacteriol. (submitted)
Moreau, S., Leret, V., Le Marrec, C., Varangot, H., Ayache, M., Bonnassie, S., Blanco, C. and Trautwetter, A. (1994) Res. Microbiol. (in press)
O'Gara and Dunican (1994) Biochem. Biophys. Res. Comm. (in press)
Oguiza, J.A., Malumbres, M., Eriani, G., Pisabarro, A., Mateos, L.M., Martin F. and Martín, J.F. (1993) J. Bacteriol. **175**:7356-7362
Peters-Wendisch, P.G., Eikmanns, B.J., Thierbach, G., Bachmann, B. and Sahm, H. (1993) FEMS Microbiol. Lett. **112**:269-274
Pisabarro, A., Malumbres, M., Mateos, L.M., Oguiza, J.A. and Martín, J.F. (1993) J. Bacteriol. **175**:2743-2749
Reinscheid, D.J., Eikmann, B.J. and Sahm, H. (1994) Microbiology (submitted)
Reinscheid, D.J., Eikmanns, B.J. and Sahm, H. (1994) J. Bacteriol. 176 (in press)
Reinscheid, D.J., Kronemeyer, W., Eggeling, L., Eikmanns, B.J. and Sahm, H. (1994) Appl. Environm. Microbiol. **60**:126-132
Schäfer, A., Kalinowski, J. and Pühler, A. (1994) Appl. Environ. Microbiol. **60**:756-759
Schäfer, A., Tauch, A. Jäger, W., Kalinowski, J., Thierbach, G. and Pühler, A. (1994) Gene **145**:69-73
Schäfer, A., Schwarzer, A., Kalinowski, J. and Pühler, A. (1994) J. Bacteriol. (in press)
Schwarzer, A. and Pühler, A. (1991) Bio/Technology **9**:84-87
Schwinde, J.W., Thum-Schmitz, N., Eikmanns, B.J. and Sahm, H. (1993) J. Bacteriol. **175**:3905-3908
Tauch, A., Kassing, F., Kalinowski, J. and Pühler, A. (1994) (submitted)
Tauch, A., Kirchner, O., Wehmeier, L., Kalinowski, J. and Pühler, A. (1994) (submitted)
Wohlleben, W., Muth, G. and Kalinowski, J. (1993) In: Biotechnology (Eds. Rehm HJ, Reed G, Pühler A, Stadler P), volume 2 'Genetic fundamentals and genetic engineering' (Ed. Pühler, A.), pp. 455-505. VCH Weinheim, New York

Valorization of non-conventional yeasts of industrial interest: exploration and molecular engineering of their genetic constituents (BIOT CT-910267)

COORDINATOR:
1. H. FUKUHARA, Institut Curie, Orsay, F
PARTICIPANTS:
2. L. FRONTALI, Univ Rome, I
3. C.P. HOLLENBERG, Univ Düsseldorf, D
4. A. DOMINGUEZ, Univ Salamanca, E
5. P.A. MEACOCK, Univ Leicester, GB
6. I. FERRERO, Univ Parma, I
7. C. GAILLARDIN, INRA, Thiverval-Grignon, F
8. R. FLEER, Rhône-Poulenc Rorer, Vitry-sur-Seine, F
9. J. VERBAKEL, Unilever, Vlaardingen, NL

BACKGROUND INFORMATION

The non-conventional yeasts, with their wide variety of metabolic capabilities, represent a large genetic resource that has so far been poorly exploited in biotechnology. Through the preceding BAP some members of this project formed a research network to initiate a study of one of these organisms, the dairy yeast *Kluyveromyces lactis*, and to examine its potential for biotechnological processes. That venture showed that the non-conventional yeasts can offer new and more efficient routes for the production of recombinant proteins of medical, industrial and agricultural importance. This programme stems from that initiative and seeks to continue research on selected species of non-conventional yeasts of proven industrial usefulness, in order to develop them fully as genetically manipulatable organisms suitable for modern biotechnology.

OBJECTIVES AND PRIMARY APPROACHES

The objectives were:

(i) to bring non-conventional yeasts to be genetically manipulatable organisms; in particular the dairy yeast *Kluyveromyces lactis* and the related species *Kluyveromyces marxianus*, the protease secreting yeast *Yarrowia lipolytica*, and the methylotrophic yeast *Hansenula polymorpha*.

(ii) to investigate the genetic basis of their unique metabolic capabilities and to develop methods for genetic manipulation of their physiology.

(iii) to create new combinations of genetic elements suitable for efficient and controlled production of useful proteins.

The approach has been to use molecular techniques to investigate aspects of genetics and physiology, particularly those that distinguish these organisms from *Saccharomyces cerevisiae*.

RESULTS AND DISCUSSION

The results are summarized and evaluated in regard to five major topics:

1. Genome Studies. (participants 1,2,3,4,5,6,7)

(i) Gene mapping in* K. lactis *chromosomes.

The *K. lactis* genome has roughly the same size as the *S. cerevisiae* genome (15 Mbp), but the number of chromosomes is fewer (6 vs. 16). Newly cloned genes, including those contributed by all the participants (See section 2), as well as those from external laboratories have been used for the construction of a physical and genetic map. Many mutations have been isolated and mapped with priority for those involved in carbon metabolism and regulation of fermentation. Several new auxotrophic mutations were also isolated, mapped and some of the relevant genes cloned. Fig.1 thus obtained is the first genetic map of the *K. lactis* genome. The map contains more than 70 genes and several genetic linkage groups. The local gene order of some genes was shown to be similar to that in *S. cerevisiae*. This raises the question of how equivalent genes are distributed in the two species whose chromosome patterns are quite different. Interestingly, the *K. lactis* genome does not seem to be as redundant as the *S. cerevisiae* genome, at least for glycolysis genes, thus facilitating mutational analysis of fermentation. *K. lactis-S. cerevisiae* shuttle libraries have been constructed and used in isolation and comparative studies of related genes in the two species (*e.g.*, genes of the galactose regulon, glucose transporters and alcohol dehydrogenases). Heterologous complementation between the two species should also help functional analysis of the new genes found by the ongoing EC programme on yeast genome sequencing (in which some of the participants are involved).

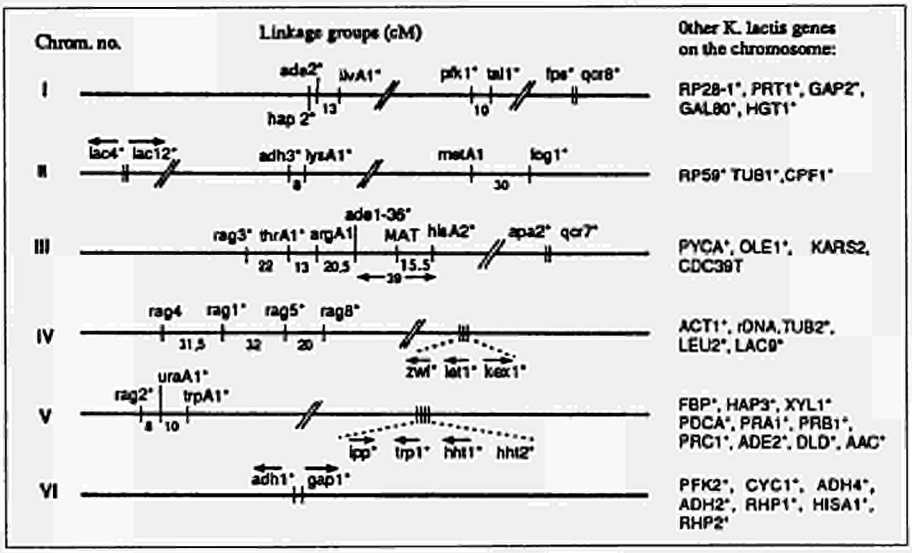


Fig. 1. Genetic map of Kluyveromyces lactis

(ii) Gene Mapping in **Yarrowia lipolytica**

The genome of this yeast is slightly larger than that of *K. lactis*. The first genetic map of *Y. lipolytica*, described more than ten years ago, contained 5 genetic linkage groups. Since then, genetic studies have been scarce. During the project, mapping studies progressed, together with karyotype analysis. A systematic comparison of various isolates revealed a notable polymorphic variation which probably explains the initial difficulty in setting up a genetic system in this yeast. The genome of laboratory strains contain five chromosomes. 33 genes (of which 28 are new) were physically mapped to individual chromosomes. At present, 59 genes are identified on the chromosomes. Three centromeres were cloned together with autonomously replicating sequences. The organization of this genome seems to be very dissimilar to those of *S. cerevisiae* and *K. lactis*, for example, we note the dispersed location of ribosomal DNA clusters on four chromosomes and the difference in gene transcription signals. Such knowledge on the genome organization is of practical importance in the heterospecific manipulation of genetic material involving different non-conventional species.

2. Gene Isolation and Characterisation. (participants 1, 2, 3, 4, 5, 6, 7, 8, 9)

A primary objective has been to identify, isolate and characterise genes and promoters useful for the biotechnological exploitation of non-conventionnal yeasts. Our aim was to have available a range of regulatable promoters suitable for manipulation of cellular metabolism and expression of heterologous genes in response to various stimuli and growth regimes. This endevour has been remarkably successful leading to the isolation of several *K. lactis* genes regulated by carbon source (*Kl.ADH1,2,3* and *4, Kl.RAG5* and *8, Kl.GAL80, Kl.DLD1*) by stress and starvation (*Kl.HSP12*), by thiamine (*Kl.THI4* and *5*, originally called *MOL1* and *MOL2*, respectively) and by phosphate (*Kl.PHO5*). In the related yeast *K. marxianus* the strongly expressed, but glucose-repressible, inulinase gene has been recovered and explored as a potential expression system (Section 5).

Of the four *K. lactis* alcohol dehydrogenase genes, only *Kl.ADH3* and *Kl.ADH4* encode products with amino-terminal extensions indicating that these proteins might have a mitochondrial location. While *Kl.ADH1* and *Kl.ADH2* are essentially constitutive the other two are regulated by carbon source; *Kl.ADH3* is induced by respiratory carbon sources, except ethanol, and repressed by glucose; *Kl.ADH4* is specifically induced by ethanol but is insensitive to glucose repression. Moreover *Kl.ADH4* is constitutively expressed in Rag^+ strains, which produce intracellular ethanol. Mutants in each of these four genes have no evident phenotype indicating that the loss of any one activity is compensated by the others. *Kl.DLD1* also encodes a product with an N-terminal extension typical of a mitochondrial-located protein and is specifically induced by lactate and repressed by glucose; the mitochondrial enzyme D-LCR is essential for oxidation of D-lactate. The *RAG* series genes which have roles in regulation of carbohydrate metabolism are described in Section 3.

The gene *Kl.HSP12* is not only induced by glucose depletion, but also by a variety of physiological stress situations that include, heat shock, osmotic shock, alcohol shock and oxidative stress; it encodes a 12kDa highly hydrophilic protein showing 60% identity with its *S. cerevisiae* counterpart.

Kl.PHO5 encodes a 465 aminoacid secreted protein homologous to the *S. cerevisiae PHO5* product. Expression studies showed that the gene is regulated by the availability of inorganic phosphate, and both a potential regulatory sequence within

the promoter and a positive regulator protein have been identified. *Kl.THI4* and *Kl.THI5* encode enzymes of the thiamine (vitamin B1) biosynthetic pathway. Both are expressed to extremely high levels in thiamine-free medium, but are completely repressed by addition of the vitamin to the growth medium. The time of their induction during the course of a batch culture can be controlled by manipulation of the level of vitamin supplied in the medium. Gene disruption experiments showed that *Kl.THI4* and *Kl.THI5* are involved in the unresolved pathways leading to the formation of the two thiamine biosynthetic precursors, hydroxyethyl-thiazole and hydroxymethyl-pyrimidine, respectively.

Although most gene cloning from *Yarrowia lipolytica* has focused on components of the secretory pathway (Section 4), two DNA segments responsible for regulating resistance to copper have been isolated. One, *Yl.CRF1*, encodes a 411 residue protein that is a transcriptional activator. The other contains a pair of divergently expressed genes, *Yl.MTP1* and *Yl.MTP2*. Both encode small proteins that are involved in determining copper resistance although *Yl.MTP2* is preferentially transcribed. Within the bidirectional promoter a regulatory sequence, causing induction by copper, has been localized. The glyceraldehyde phosphate dehydrogenase gene promoter has been recovered from *Hansenula polymorpha* offering a constitutive gene expression system for the methylotrophic yeast. Additionally the *Hp.MOX1* promoter has been found to experience glucose repression in *S. cerevisiae*, via the action of the *ADR1* product, indicating a conserved regulation for genes encoding peroxisomal proteins.

Our approaches to identify and isolate these genes have included use of cloned *S. cerevisiae* genes as hybridisation probes to screen genomic DNA libraries, genetic complementation of mutations in either *S. cerevisiae* or the non-conventional yeast, and direct isolation of the gene product and design of synthetic oligonucleotide hybridisation probes from the protein amino-acid sequence. All the genes cloned have been sequenced and studied at the level of transcriptional regulation. These data add substantially to knowledge of genome organisation, gene structure, codon usage, regulation, etc. and the DNA sequences provide useful information about the evolutionary relationships between yeasts. In most cases the gene promoter regions have been recovered and shown to be capable of conferring regulated high level expression of foreign DNA sequences (Section 5), by coupling to a suitable reporter gene. Thus expression systems, responding to a variety of external stimuli are now available.

Finally, studies have not been solely restricted to nuclear genes. Previous work by participant 1 on the mitochondrial genome of *K. lactis* has been extended by participant 2 in a study of transcription and processing of transcripts of a group of tRNA genes which are maintained in the same order in the mitochondrial genomes of *K. lactis, S. cerevisiae* and *Candida glabrata*. Ordered processing of polygenic transcripts has been demonstrated and intergenic processing sites have been found between the same tRNA sequences, suggesting a physiological significance in the conservation of gene order.

3. Regulation of Carbon Metabolism. (participants 1, 3, 6)

Studies on non-conventional yeasts have established the exceptional nature of *S. cerevisiae* in the regulation of carbon metabolism. No other yeast species studied so far showed a similar preference for fermentative metabolism under aerobic growth conditions and comparably effective mechanisms of glucose repression. A great deal about the molecular basis for such differences in physiology between

yeast genera has been learnt from comparative studies in *K. lactis* which suggested that the regulation of sugar uptake may be a major source of variations.

Screening of mutants that either negatively affect fermentative metabolism (Rag$^-$ mutants) or affect carbon catabolite repression (*dgr, fog, hxt* mutants) have converged and created a link between the efficiency of fermentation and sugar dependent signalling processes like glucose induction or glucose repression. Several genes complementing these mutants have been cloned and sequenced (*RAG1, RAG2, RAG3, RAG4, RAG5, RAG8, FOG1, HXTB2*). Interaction between these genes have revealed a regulatory network between the rate of sugar uptake, a regulatory function of hexokinase, protein phosphorylation and transcriptional regulation by glucose.

The important role of sugar uptake kinetics was seen by the analysis of the *RAG1* gene which encodes an inducible low-affinity sugar carrier of the highly conserved *HXT* multigene family of facilitative transporters. In contrast to *S. cerevisiae*, where seven members of this family have been isolated, there is little redundancy in *K. lactis*. In a few strains a second gene *HXTB2* with 85% similarity is located next to *RAG1* and can complement for its function. The *RAG1* gene was isolated twice, by complementation of the Rag$^-$ phenotype (selecting for resistance against antimycin A on glucose) and by its ability to establish glucose repression of the β-galactosidase gene. Glucose repression of several other genes was also affected in a *rag1* mutant and induction by glucose of the pyruvate decarboxylase gene is impared.

Another permease gene, *HGT1*, has also been isolated. *HGT1* seems to be a unique gene encoding a high-affinity glucose carrier but it is not present in all strains.
Three genes, *RAG4, RAG5* and *RAG8*, were found to be required for the glucose inducibility of *RAG1*. *RAG4* and *RAG5,* but not *RAG8*, also interact with *HGT1*: *RAG5* positively controls the expression of *HGT1*, and *RAG4* has a negative effect. *RAG5* encodes the only hexokinase gene in *K. lactis*. Its activity is lacking in two classes of catabolite non-repressible 2-deoxy-D-glucose resistant mutants, DGR 148 and DGR 239, suggesting a requirement for those gene products. The *rag5* mutants abolish glucose repression of several genes and glucose induction of *RAG1* and the pyruvate decarboxylase (PDC) gene. This establishes a regulatory role for hexokinase in sugar metabolism similar to the one of hexokinases (PII encoded by *HXK2*) in *S. cerevisiae*. Mutants lacking hexokinase have been identified both as rag5 and as 2-deoxyglucose resistant mutant. Heterospecific complementation experiments showed that *RAG5* of *K. lactis* and *HXK2* of *S. cerevisiae* are interchangeable for the catalytic activity of their products but not for their role in glucose repression, indicating that sugar phosphorylation and the regulatory function are separable activities of hexokinase and that the regulatory function may require specific protein-protein interactions.

Another member in the regulatory cascade is *RAG8*, which shares 65% identity with two casein kinase type I genes *YCK1* and *YCK2* of *S. cerevisiae*. The transcription of *RAG8* is itself inducible by glucose.

RAG3 is a gene that specifically affects PDC induction by glucose. Cloning and sequencing showed that the deduced *RAG3* protein sequence shares identity with the *PDC2* product of *S. cerevisiae* which is a regulator of pyruvate dehydrogenase. *RAG3* is poorly transcribed, not regulated by carbon source and affects expression of PDC structural gene at the transcriptional level.

Two genes *FOG1* and *FOG2* have been isolated that affect the kinetics of derepression from catabolite repression of maltase, L-lacate dehydrogenase, β-galactosidase and invertase. *FOG2* also seems to be required for sporulation. *FOG1* has been sequenced. It showed 71% and 62% similarity with *SPM1* and *SPM2* (dominant suppressors of RNA polymerase mutations) of *S. cerevisiae*, respectively. The gene is constitutively expressed at a low level.

Studies on the regulation of lactose and galactose metabolism have progressed considerably, and now provide a rather detailed insight into the regulatory mechanisms at the molecular level. Regulation depends on the interaction between the transcriptional activator Lac9, a *K. lactis* homologue of Gal4 and its inhibitor Gal80. The *Kl.GAL80* gene was cloned and shown to be highly inducible by Lac9. During induction, Gal80 repressor is inactivated by a regulatory function of galactokinase which is regulated by the intracellular concentration of galactose. Thus again the rate of sugar uptake and the sugar kinase play key roles in the signalling process leading to induction. Since permease and kinase, being activators of Lac9 function, are themselves encoded by Lac9-controlled genes, a feedback loop is formed implying that the Lac9 protein controls its own activity.

4. Genetic Analysis of the Secretion Apparatus. (participants 4 , 7)

Several examples have shown that non-conventional yeasts are often superior to *S. cerevisiae* for the secretion of foreign proteins. In a few cases however, disappointing results were obtained, but not further analyzed because of lack of basic knowledge and mutants of the secretory apparatus in these yeasts. A primary goal was to identify and characterize genes of the secretory apparatus which might be useful to optimize secretion of heterologous proteins. This work was particularly developed on *Yarrowia lipolytica* strains, which are natural secretors of high molecular weight polypeptides into the growth medium.

Several genes encoding components of the endoplasmic reticulum (ER)-translocation machinery have been isolated and characterized. Homologs of *SEC61* and *SEC62*, which form part of the translocon in the ER membrane, have been isolated either by PCR on genomic DNA, or by functional complementation of *S. cerevisiae* mutants using a *Y. lipolytica* cDNA library made in a *S. cerevisiae* expression vector. Both cDNA and genomic clones were analyzed. Conservation of aminoacid sequence is higher for *SEC61* (68% identity) than for *SEC62* (35%). Disruption of these genes is in progress, as well as amplification to see if they may suppress defects in the early steps of protein secretion (see co-lethal mutations and suppressors of SRP mutations below). The *SEC65* gene was isolated fortuitously as a gene adjacent to *URA5*. *SEC65* is an essential gene and encodes a protein homologous to canine, human and *S. cerevisiae* Srp19 (one of the polypeptides associated to the signal recognition particle SRP).

In an effort to isolate genes affecting late steps of the secretory pathway, we tried to isolate *SEC4* homologs encoding small GTP-binding proteins, of the Rab family. A gene library of *Y. lipolytica* was screened with an oligonucleotide deduced from a highly conserved sequence in the Rab family. Four genes were isolated: *RYL1* was tightly related to *SEC4*, *RYL2* was loosely related to *RAB2*, *RYL3* to *CDC42* (controlling cell polarity) and RYL4 to the *RHO* subfamily. *RYL1* was shown to be an essential gene and, when put under the control of the *S. cerevisiae GAL10* promoter, to relieve specifically the thermosensitivity of the *sec4-8* mutant of *S. cerevisiae*. It thus seems that *RYL1* is a functional homolog of *SEC4* involved in the fusion of secretory vesicles with the plasma membrane. *RYL2* seems to identify a new member of the YPT/rab family, whose function remains to be assessed.

Two different approaches were developed in order to identify new genes involved in ER translocation:

(i) Suppresors of a 7S RNA thermosensitive mutant (Scr 2-II-13, affecting signal recognition particle docking on the ER or release from the ribosome) were identified. Five such suppressor genes which retain a thermosensitive phenotype in the presence of a wild type copy of *SCR2* were saved for further analysis. One of them (*TSR5*) is itself suppressed by a multicopy vector carrying *SEC65*. Another one (*TSR1*) identifies a new gene. Genes complementing *TSR1* were isolated from a replicative library made with wild type DNA and are currently still being analyzed.

(ii) Mutations enhancing the secretory defect of Scr 2-II-13 were identified as co-lethal events. One of these results in a thermosensitive secretory defect in the presence of a wild type *SCR2* allele. The coresponding gene was cloned (226), and shown to encode a 420 aminoacid polypeptide with no homolog in data banks. The 226p is predicted to carry an N-terminal signal sequence and a C-terminal ER-retention signal. Antibodies against a GST-fusion of 226p were raised and evidenced an abundant protein localized in a membrane-bound compartment. No regulation of 226p synthesis could be detected (heat shock, accumulation of non-secretable mutant proteins). Overexpressing the wild type 226p severely limits transit of secretory proteins prone to misfolding (glycosylation mutant) as well as transit of wild type AEP (alkaline exocellular protease). Deletion of 226 results in a thermosensitive phenotype associated with a drastic reduction of secretory protein synthesis. Preliminary results of immunoprecipitation after cross-linking cell extracts suggest that 226p may be associated to proAEP. These results suggest that 226p may bind nascent secretory proteins to couple translocation into the ER to proper folding and exit from the translocation apparatus. Three other genes were isolated by complementation of two independently isolated co-lethal mutations and are currently under analyis.

The *SEC14* gene has been cloned in a collaboration between participants 4 and 7. It was shown that Sec14p is a Golgi associated protein by immunofluorescence labelling, and that it was endowed with a phosphatidylcholine/ phosphatidylinositol (PI/PC) transfer activity, as already known for its *S. cerevisiae* homolog. Contrary to the latter however, *SEC14* is not an essential gene in *Y. lipolytica* and its deletion does not detectably affect protein secretion in this yeast. Strains deleted for the *SEC14* gene were however unable to execute the yeast-mycellium transition characteristic of this dimorphic yeast. These results point to an unsuspected role of PI/PC transfer proteins in cell differentiation.

5. Heterologous Expression and Fermentation Studies.
(participants 1, 2, 3, 4, 5, 7, 8, 9)

The experimental work with regard to this topic was focused on

(i) the development of new or improved cloning and expression vectors for *Kluyveromyces* and *Yarrowia*,
(ii) the study of heterolgous gene expression driven by new transcriptional promoters, and
(iii) the study of optimized culture conditions in batch, fed-batch and chemostat fermentations.

The heterologous proteins whose expression was examined comprised bacterial ß-glucuronidase (*GUS), plant* α-galactosidase (α-GAL), human interleukin-1β (IL-1β) and human serum albumin (HSA). Except for *GUS* which is an intracellular enzyme, all of these proteins were secreted to the culture supernatant.

(i) Host/vector systems. In addition to the now well-characterized 2 μ-like multi-copy vectors based on the circular plasmid pKD1, a number of additional tools have been developed for *K. lactis*. These include low-copy number centromeric vectors (CEN2/KARS12) and integrative vectors that can be targeted to the *LAC4* locus by fluoro-orotic acid resistance (FOA^r) selection in specialized *lac4::URA3* host strains such as JA6/D4. Similar levels of *GUS* expression (70-100 mU/mg) were obtained whether the bacterial gene was integrated at the chromosomal *LAC4* locus or carried, under *LAC4* promoter control, on a centromeric vector. The effect of gene dosage has been studied with regard to HSA secretion. Expression cassettes carrying the prepro-HSA cDNA gene under control of various constitutive and regulated promoters (*LAC4, Sc.PGK, Sc.GPD, Kl.ADH3, Kl.ADH4*) were either chromosomally integrated (1 to 2 copies) or engineered into pKD1-derived vectors. In all cases studied, at least 10 to 20-fold higher product levels were obtained with the pKD1 system, indicating that, at least with regard to HSA, gene dosage is a crucial parameter for optimal product yield even when using strong transcriptional promoters. Regulated expression under control of the *LAC4* or *Kl.ADH4* promoter was maintained when using multi-copy vectors.

A new native plasmid that is structurally related to pKD1, without, however, exhibiting significant sequence homology with the latter, has been identified in a strain of *K. waltii* and used to engineer shuttle vectors able to replicate in *E. coli* as well as in several *Kluyveromyces* species. Interestingly, pKW1 and pKD1 differ in their host range. Expression vectors based on pKW1 have been used to achieve high level secretion of IL-1β from *K. waltii*. Another new type of expression vector is based on the linear, cytoplasmically located killer-plasmid pGKL1 of *K. lactis*. Due to the presence of proteins covalently attached to the termini, the *in vitro* modification of these linear plasmids is precluded. We therefore engineered vectors that target the gene of interest to pGKL1 by homologous recombination *in vivo*. Expression of the IL-1β cDNA gene driven by a pGKL1 promoter could be obtained, albeit at rather low levels. A mutant cytoplasmic promoter with higher transcriptional activity could be isolated and is currently under study.

The analysis of three chromosomal DNA fragments (KARS12, KARS20 and KRF4) functioning as autonomously replicating sequences (ARS) in *K. lactis* has led to the identification of the regions (100-300 bp) that are crucial for ARS function and binding of ABF1 protein. KARS20 and KRF4, but not KARS12, confer autonomous replication to plasmids both in *K. lactis* and *S. cerevisiae.*

As far as *Y. lipolytica* is concerned, a series of vectors allowing multi-copy integration at the ribosomal DNA (rDNA) loci have been constructed, using a promoter-truncated version of the *URA3* gene as selectable marker. Depending on the vector used, up to 60 copies of a secreted protease encoding gene could be integrated into the genome. However, whereas the copy number remained stable under conditions where the reporter gene was not expressed, growth arrest and rapid selection of transformants with lower copy number were observed upon induction of the reporter gene. A similar result was obtained when targeting the plant α-galactosidase gene under control of the inulinase promoter to the rDNA locus of *K. marxianus.*

(ii) Promoter studies. The cloning of a wide variety of new *K. lactis, K. marxianus, H. polymorpha* and *Y. lipolytica* genes in the course of the present study has provided new transcriptional signals for the expression of heterologous genes. A number of these promoters have been fused to foreign genes and their expression has been studied in their respective hosts. The bacterial *GUS* reporter gene has been expressed in *K. lactis* under control of the *Kl.ADH4, Kl.LAC4, Kl.GAL80, Kl.THI5* and *Kl.HSP12* promoters. The *Kl.ADH4* promoter has also been studied

for expression of the bacterial kanamycin resistance gene and the human serum albumin gene. The plant α-galactosidase gene has been expressed under control of the *Kl.PHO5, Sc.PGK* and *Sc.GAL7* promoters in *K. lactis*, the *Yl.MTP1* promoter in *Y. lipolytica*, and the inulinase promoter in *K. marxianus*. In all cases studied, transcriptional regulation followed that of the native promoter context (Section 2).

(iii) Fermentation studies. Batch, fed-batch and chemostat studies have been carried out with a *K. lactis Δpgk* strain transformed with a pKD1-based plasmid carrying the HSA gene under control of the *Kl.ADH4* promoter and the intact *Kl.PGK* gene as a selectable marker. This host/vector system was found to be stable over at least 200 generations when grown in the presence of fermentable carbon sources. In batch cultures, the secretion of HSA reached up to 300 mg/liter, with a yield of 10-15 mg/g of carbon substrate. Chemostat studies allowed the optimization of culture conditions, which were then applied to high cell density fermentations. Fed-batch conditions were achieved either by coupling the carbon dioxide evolution rate (CER) to the nutrient feed-rate, or, after initial batchwise growth and consumption of the C-substrate, by adding the feed medium according to a pre-programmed time-based feeding profile. Growth as well as production kinetics were found to be well correlated to the amount of substrate added. Using a completely defined medium, HSA production reached 840 mg/liter at a biomass of 69 g/liter (dry cell weight) and a yield (Yx/s) of 45%. With a complex medium, product levels (secreted, correctly processed HSA) were above 1 g/liter.

Batch and fed-batch studies confirmed that the presence of an additional chromosomal *LAC9/KlGAL4* allele (Section 3) significantly increases the expression levels of homologous (yeast β-galactosidase) and heterologous (bacterial β-glucuronidase) genes placed under control of the *K. lactis LAC4* promoter. However, higher *LAC9* gene dosage led to a loss of glucose repression otherwise observed with the specific strain used in this study (JA6).

The observation that *K. marxianus* is able to secrete high levels of the enzyme inulinase stimulated efforts to optimize fermentation conditions and apply these to the expression of the enzyme α-galactosidase. Fed-batch propagation of *K. marxianus* strain CBS6556 under RQ control resulted in the production of >2 g/liter of inulinase within 45 hours. Unfortunately, α-GAL secretion under control of the inulinase promoter and prepro-sequence was at least 10-fold lower, whether the plant gene was expressed from a pKD1-based vector or integrated at chromosomal locations. Northern blot analysis showed that α-GAL specific mRNA was present in high amounts, indicating a post-transcriptional bottleneck, possibly within the secretory pathway.

6. Conclusion

This project has convincingly shown that non-conventional yeasts have considerable potential as efficient systems for the production of recombinant proteins. A number of systems for the genetic and molecular manipulation of these organisms have been developed and successfully applied to fermentation studies in which foreign proteins have been produced at high levels. The fundamental knowledge gained about the genetics and physiology of these yeasts, particularly in relation to their control of carbon metabolism, protein secretion and gene regulation, constitutes an important contribution to the subjects of Genetics and Microbiology. Overall, the knowledge and materials that this study has generated provide a solid basis for the rational biotechnological exploitation of these important natural resources in the future.

MAJOR SCIENTIFIC BREAKTHROUGHS

An effort was made to assimilate all aspects of knowledge about these organisms. Thus the first review of *K. lactis* biology has been produced during the project, and the first detailed genetic map constructed. This has been circulated widely within the international *Kluyveromyces* community. Research undertaken on *K. lactis* has provided important new information about the regulation of carbon metabolism in yeasts, and clarified understanding of fermentative metabolism in *S. cerevisiae*. New regulated genes have been isolated and characterised from all of the yeasts studied, and a number of have been used successfully for controlled production of foreign proteins. Some gene promoters have been used to construct expression systems giving high yields of heterologous proteins in industrial fermentations. High level production of inulinase from *K. marxianus* has been achieved, but it is clear that more research is needed in order to fully develop this species for production of foreign proteins. Studies on *Y. lipolytica*, a species noted for its high capacity of protein secretion, have led to the isolation and characterisation of several genes of the secretory pathway, allowing its manipulation and comparison with other less efficient organisms.

MAJOR COOPERATIVE LINKS

(i) Coordination meetings and collective interactions:

The annual European workshop 'Biology of *Kluyveromyces*', created in 1988 during BAP, continues its regular activity with steadily increasing numbers of participants. These meetings have been efficient instruments for the coordination of activities of the consortium and for its communication with a wider community. The 4th *Kluyveromyces* workshop was organized in Düsseldorf by participant 3 (1991); the fifth in Vienna by participants 1 and 2 (1992), the 6th, in Siena by participant 6 (1993) at the site of the company IRIS. The 7th meeting is planned by participant 4 in Salamanca (1994). The annual coordination meetings of the consortium were held at Orsay, Düsseldorf and Vienna.

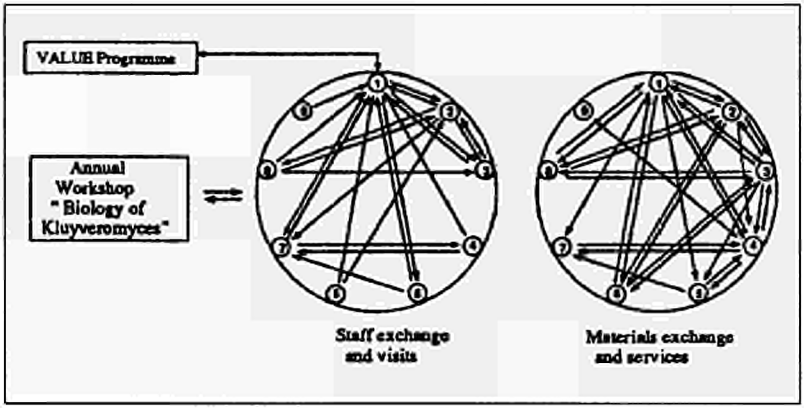


Fig. 2. Cooperative links

(ii) Individual interactions:

The consortium had multilateral cooperative activities. Interactions involved exchanges of staff, material, training and teaching services. Cooperation with industry is illustrated by the transfer of the new regulatory systems studied by

participants 2 and 3 to participant 8 for fermentation studies; a joint patent has been applied by participants 1 and 8. The interactive links are summarized in Fig.2. (in which an arrow generally corresponded to several actions).

PUBLICATIONS

Joint publications

Lopez, M.C., Nicaud, J.M., Skinner, H.B., Vergnolle, C., Kader, J.C., Bankaitis, V.A. and Gaillardin, C.J. A phosphatidylinositol/phosphatidylcholine transfer protein is required for differentiation of the dimorphic yeast *Y. lipolytica* from the yeast to the mycellial form. Cell Biol.124, 113-127 (1994). (Participants 4 and 7).

Sanchez, M., Beckerich, J.M., Gaillardin, C. and Dominguez, A. Isolation and cloning of *Yarrowia lipolytica SEC65* gene, a component of the yeast signal recognition particle with homology with the human *SRP19* gene. (submitted). (Participants 4 and 7).

Dominguez, A. and Gaillardin, C. Codon usage in the yeast *Yarrowia lipolytica*. (submitted). (Participants 4 and 7).

Wésolowski-Louvel, M., Breunig, K.D. and Fukuhara, H. *Kluyveromyces lactis*. In: Wolf, K. (ed.), Genetics, Biochemistry and Molecular Biology of Non-conventional Yeasts, a Laboratory Manual. Springer-Verlag, Berlin. (in press). (Participants 1 and 3).

Jung, G., Saliola, M., Yeh, P., Fournier, A., Mazzoni, C., Crisa, A., Falcone, C. and Fleer, R.Utilization of the *KlADH4* promoter for the regulated expression of heterologous genes in the yeast *Kluyveromyces lactis*. 6th European Congress on Biotechnology, Florence, June 13-17, 1993. (Participants 2 and 8).

Ménart, S., Wésolowski-Louvel, M., Fleer, R., Bolotin-Fukuhara, M.Molecular cloning and partial characterization of a gene encoding a putative high-affinity glucose transporter of *Kluyveromyces lactis*. Workshop 'Biology of *Kluyveromyces* VII', Siena, June 18-19,1993. (Participants 1 and 8).

Goffrini, P., Ferrero, I. and Breunig, K.D. Characterization of carbon catabolite repression in *rag* mutants of *Kluyveromyces lactis*. in preparation. (Participants 3 and 6).

Zenke, F.T., Goffrini, P., Breunig, K.D. and Ferrero, I. Differentiation between glucose and galactose repression in *Kluyveromyces lactis*. (in preparation). (Participants 3 and 6).

Jung, G., Fleer, R., Saliola, M. and Falcone, C. Use of the *KlADH4* promoter for the ethanol-dependent production of heterologous proteins in *K. lactis*. (in preparation).

Individual publications

Zhang, Y.P., Chen, X.J., Li, Y.Y. and Fukuhara, H. *LEU2* gene homologue in *Kluyveromyces lactis*. Yeast 8, 801-804 (1992)

Sor, F. Electrophoretic technique for chromosome separation. Methods in Molec. Cell. Biol. 3,65-70 (1992)

Chen, X.J., Cong, Y.S., Wésolowski-Louvel, M., Li, Y.Y. and Fukuhara, H. J. Characterization of a circular plasmid from the yeast *Kluyveromyces waltii*. Gen. Microbiol.138, 337-345 (1992)

Prior, C., Mamessier, P., Fukuhara, H., Chen, X.J. and Wésolowski-Louvel, M. The hexokinase gene is required for transcriptional regulation of the glucose transporter gene *RAG1* in *Kluyveromyces lactis*; Mol. Cell. Biol. 13, 3882-3889 (1993)

Prior, C., Fukuhara, H., Blaisonneau, J. and Wésolowski-Louvel, M. Low-affinity glucose carrier gene *LGT1* of *Saccharomyces cerevisiae*, a homologue of the *Kluyveromyces lactis RAG1* gene. Yeast 9, 1373-1377 (1993)

Sanchez, M., Iglesias, F.J., Santamaria, C. and Dominguez, A. Transformation of *Kluyveromyces lactis* by electroporation. Appl. Env. Microbiol. 59, 2087-2092 (1993)

Dominguez, A. Yeast metallothioneins, gene amplification and biotechnological application. pp. 183-198. *In:* Profiles in Biotechnology' (Villa,T.G. and Abalde, J. eds.), University of Santiago. 1992.

Ragnini, A. and Frontali, L. Ordered processing of the polygenic transcripts from a mitochondrial tRNA gene cluster in *K. lactis*. Curr. Genet. 25, 342-349 (1994)

Zachariae, W., Kuger, P. and Breunig, K.D. Glucose repression of lactose/galactose metabolism in *K. lactis* is determined by the concentration of the transcriptional activator LAC9 (KlGAL4). Nucl. Acids Res. 21, 69-77 (1992)

Zachariae, W. and Breunig, K.D. Expression of the transcriptional activator LAC9 (KlGAL4) in *Kluyveromyces lactis* is controlled by autoregulation. Mol. Cell. Biol. 13, 3058-3066 (1993).

Zenke, F., Zachariae, W., Lunkes, A. and Breunig, K.D. GAL80 proteins of *Kluyveromyces lactis* and *Saccharomyces ceresisiae* are highly conserved but contribute differently to glucose repression of the galactose regulon. Mol. Cell. Biol. 13, 7566-7576 (1993).
Naumov, G., Naumov, E., Fournier, P., Nguyen, V. and Gaillardin, C. Chromosomal polymorphism in *Y. lipolytica* and related species. Curr. Genet. 23, 450-454 (1993).
Walsh, D.J. and Meacock, P.A. Stress responsive / growth phase dependent genes of *Kluyveromyces lactis*. Proceedings of the Workshop 'Biology of *Kluyveromyces*. V', Vienna,1992.
Fleer, R. Engineering yeast for high level expression. Curr. Opinion Biotechnol. 3, 486-496 (1992).
Gonzalez, F.J., Fauste, C., Burguillo, F.J. and Dominguez, A. Kinetic behaviour of a repressible acid phosphatase from the yeast *Yarrowia lipolytica*: A comparative study between the solubilized enzyme, the enzyme bound to cell wall fragments and the enzyme bound to intact cells. Biochim. Biophys. Acta. 1162, 17-27 (1993).
Saliola, M., Bellardi, S., Marta, I and Falcone, C. Ethanol metabolism in the yeast *Kluyveromyces lactis*. Proceedings of the 6th European Congress of Biotechnology (ECB6), edited by Alberghina, L., Frontali, L. and Sensi, P. Progress in Biotechnology 9, 317-321 (1994)
Saliola, M., Bellardi, S, Marta, I. and Falcone, C. Glucose metabolism and ethanol production in *adh* multiple and null mutants of *Kluyveromyces lactis*. Yeast, (in press).
Walsh, D.J. and Meacock, P.A., Thiamine repressible genes of *Kluyveromyces lactis*. Proceedings of the Workshop 'Biology of *Kluyveromyces*. VI', Siena, 1993.
Cong. Y.S., Yarrow, D., Li, Y.Y. and Fukuhara, H.J. Linear DNA plasmids from the yeasts, *Pichia etchellsii, Debaryomyces hansenii* and *Wingei robertsiae*. Microbiol. 140, 1327-1335 (1994).
Cong, Y.S., Wésolowski-Louvel, M. and Fukuhara, Creation of a functional promoter by rearrangement in a *Kluyveromyces lactis* linear plasmid. H. Gene. **147**, 125-129 (1994).
Le Dall, M.T., Nicaud, J.M. and Gaillardin, C. Multiple copy integration in the yeast *Yarrowia lipolytica*. Curr. Genet. **26**, 38-44 (1994)
Luani, D., Lodi, T. and Ferrero, I. Genes coding for mitochondrial proteins are more strongly biased in *Kluyveromyces lactis* than in *Saccharomyces cerevisiae*. Curr. Genet. **26**, 919 (1994)
Lodi, T., O'Conner, D., Goffrini, P. and Ferrero, I. Carbon catabolite repression in *Kluyveromyces lactis*: isolation and characterization of the KlDLD gene encoding the mitochondrial enzyme D-lactate ferricytochrome c oxidoreductase. Mol. Gen. Genet. **244**, 622-629 (1994).
Weirich, J., Kuger, P., Hasse, V. and Breunig, K.D. The influence of glucose transporters on glucose signalling in the yeast *Kluyveromyces lactis*. (submitted).
Sanchez, M., Prado, M., Iglesias, F.J. and Dominguez, A. Cloning and sequencing of the *URA5* gene in the yeast *Yarrowia lipolytica*. (submitted).
Ferminan, E. and Dominguez, A. Isolation and analysis of the *Kluyveromyces lactis PHO5* gene, encoding a secreted repressible acid phosphatase. Purification and properties of the enzyme. (submitted).
Zachariae, W. and Breunig, K.D. A feedback control mechanism for the transcriptional activator Gal4/Lac9. (submitted).
NGuyen, C., Bolotin-Fukuhara, M., Wésolowski-Louvel, M. and Fukuhara, H. The respiratory system of *Kluyveromyces lactis* escapes from the HAP2 control. Gene, (in press).
Prado, M., Gonzalez, F.J. and Dominguez, A. Isolation and sequence analysis of metallothionein-encoding genes in the yeast *Yarrowia lipolytica*. (submitted).
Prado, M., Gonzalez, F.J. and Dominguez, A. Cloning and structural characterization of *CRF1* gene required for copper resistance in the yeast *Yarrowia lipolytica*. (submitted).
Goffrini, P., Ficarelli, A. and Ferrero, I. Carbon catabolite repression in *Kluyveromyces lactis*: a comparison with *Saccharomyces cerevisiae*. (submitted).
Viola, A.M., Galeotti, C.L., Goffrini, P., Ficarelli, A. and Ferrero, I. A *Kluyveromyces lactis* gene homologue to *AAC2* complements the *Saccharomyces cerevisiae op1* mutation. (submitted).
Pertuiset, B., Beckerich, J.M. and Gaillardin, C. Cloning of *SEC4* homologue from the yeast *Yarrowia lipolytica*. (submitted).
Saliola, M. and Falcone, C. Regulation by carbon source of the transcription of *ADH* genes in *K. lactis*. (in preparation).

Fabiani, L., Frontali, L. and Newlon, C. A *Kluyveromyces lactis* ARS element requires an essential core region and auxiliary sequences for efficient function. (in preparation).
Fabiani, L. and Newlon, C. Identification of replication origin in the pKD1 plasmid by two-dimentional electrophoresis. (in preparation).
Prior, C., Fukuhara, H. and Wésolowski-Louvel, M. The *RAG3* gene of *Kluyveromyces lactis*, a homologue of the *Saccharomyces cerevisiae PDC2* gene, is involved in transcriptional regulation of pyruvate decarboxylase gene. (in preparation).
Wésolowski-Louvel, M. and Fukuhara, H. The hexokinase gene of *Kluyveromyces lactis* is involved in glucose repression. (in preparation).
Ménart, S., Billard, P., Blaisonneau, J., Bolotin-Fukuhara, M., Fleer, R., Fukuhara, H. and Wésolowski-Louvel, M. Identification and control of the expression of the high-affinity glucose transporter gene *HGT1* of *Kluyveromyces lactis*. (in preparation).
Goffrini, P., Donnini, C., Rossi C. and Ferrero, I. The *DGR1* gene is required for high-affinity glucose transport in *Kluyveromyces lactis*. (in preparation).
Goffrini, P., Ficarelli, A. and Ferrero, I. *FOG1* and *FOG2* genes are required for the derepression process in *Kluyveromyces lactis*. (in preparation).
Ficarelli, A, Goffrini P. and Ferrero, I. Isolation and molecular characterization of the *FOG1* gene of *Kluyveromyces lactis*. (in preparation).
Walsh, D.J. and Meacock, P.A. Characterization of two thiamine biosynthetic genes of the dairy yeast *Kluyveromyces lactis*. (in preparation).
Walsh, D.J. and Meacock, P.A. Isolation and characterization of a gene encoding a small stress-response protein of the dairy yeast *Kluyveromyces lactis*. (in preparation).
Walsh, D.J. and Meacock, P.A. Development of regulatable vectors for heterologous gene expression in *K. lactis* from promoters of thiamine-biosynthetic and heat shock genes. (in preparation).

Joint patent
European patent application, no. 92401711, 19 June 1992. (Participants 1 and 8).

AREA C:

CELLULAR BIOLOGY

BASIC BIOTECHNOLOGY OF PLANTS AND ASSOCIATED MICROORGANISMS

Molecular and genetic analysis of genes controlling flower development (BIOT CT-900171)

COORDINATOR:
E. COEN John Innes Institute, Norwich, GB
PARTICIPANTS:
Z. SCHWARZ-SOMMER, Max-Planck-Institut, Köln, D
J. BELTRAN PORTER, Consejo Super. de Invest. Cientificas, Valencia, ES
J. ALMEIDA, Inst. Superior de Agronomia, Lisboa, P
J. NEWBURY, Birmingham University, Birmingham, GB

BACKGROUND INFORMATION

The form, physiology and function of the flower is of central interest to plant developmental and reproductive biologists. Furthermore, the flower is of major agronomic importance both for the efficient breeding of crops and because many plant products are derived from flower seeds and fruits. European scientists have made major advances in the molecular and genetic analysis of flower development through studies on *Antirrhinum*. By studying mutations that cause *Antirrhinum* flowers to develop abnormally, they have been able to start unravelling how genes control the development of normal flowers.

OBJECTIVES AND PRIMARY APPROACHES

The aim is to isolate and study key genes controlling floral morphogenesis in an experimentally suitable system and eventually to compare them with their counterparts in other species. A range of molecular and genetic tools will be established in *Antirrhinum* to facilitate gene cloning and analysis. These include: isolation of a set of active transposons by trapping them in previously characterised genes; construction of a combined restriction fragment length polymorphism and genetic map; genetic characterisation of floral mutants, construction of double mutants and development of a transformation system. These resources will then be exploited to isolate genes by transposon-tagging. In addition, the possibility of using genes from *Antirrhinum* to study their counterparts in other species such as *Pisum sativum* will be investigated.

RESULTS AND DISCUSSION

1. Transposon-trapping

One feature of the *Antirrhinum* system, of great importance for gene isolation, is the availability of cloned transposons. If a mutation in a gene is caused by insertion of a well characterised transposon, the affected gene can usually be isolated. It is therefore essential to isolate and characterise all active transposons that are present in *Antirrhinum* lines routinely used for transposon mutagenesis. The approach adopted in this programme is to trap the active transposons in previously isolated genes. For example, the *incolorata (inc)* gene is required for red pigmentation of *Antirrhinum* flowers; *inc* mutants have pale or white flowers. 15 different *inc* mutants have been recovered by the Norwich group by screening mutagenic lines for pale or white-flowered plants. Because the *inc* gene has been cloned, it was possible to analyse the *inc* mutations at the molecular level. They showed that all of the mutants were caused by insertions of members of the CACTA family of transposons at different positions in the *inc* gene. 14 of the transposons have been

identified: Tam1 has been trapped twice, Tam2 once, Tam4 three times, Tam5 six times, Tam6 twice. The remaining transposon appears to represent a novel type and has been cloned for sequencing. In addition to *inc*, many other loci are also being used as traps in this way: five transposons have been trapped at the *plena* locus, two of which are Tam3 and one is Tam7 (Norwich), three at *squamosa* (Köln), three at *globosa*, one Tam1, one Tam7, one Tam9 (Köln) two at *deficiens* one Tam7, one Tam8 (Köln), two at *fimbriata*, both Tam3 (Norwich), two at *floricaula*, one Tam3, one CACTA type (Norwich) and five at *cycloidea*, all CACTA. The fact that the same types of transposon are repeatedly being recovered at several loci suggests that we are near saturation and have established an almost complete 'library' of transposons available for general use. In addition, because many of the trapped transposons show DNA sequence homology at their ends, it has been possible to devise a general PCR strategy for selecting for insertions in genes being investigated (Norwich). This has already enabled several new mutations to be isolated at the *fimbriata* locus.

2. Targetted tagging

Many of the mutants in *Antirrhinum* have arisen from general transposon-mutagenesis programmes. An alternative approach is to try and target a specific transposon to the gene of interest. This may be possible because several plant transposons have been shown to preferentially transpose to chromosomal positions closely linked to the site of excision. To test the feasibility of this approach, the Lisbon group carried out crosses to try and select for events in which the transposon Tam3 has moved from the pigmentation gene *pallida (pal)* to *divaricata (div)*, a locus closely linked to *pal*, which controls flower shape. About 3000 progeny from a cross between *div* and 100 different Pal^+ revertants were screened for new *div* mutants. Many of the revertants should carry copies of Tam3 in the vicinity of *div* and by growing them at 15°C, Tam3 transpositions should be favoured. Most of the progeny had flowers with a shape intermediate between *div* and wild-type, as expected because *div* is semidominant. Although no mutants with a fully *div* phenotype were observed, the experiment did provide several other interesting types of mutation.

4 plants had flowers with a near-normal shape. Molecular analysis revealed that each of these 4 plants carried an extra copy of a locus closely linked to Div^+ (< 1cM). In one of these plants it was possible to show that the previously mapped chromosomal region *pal-div-flo* had been duplicated. In the progeny derived by selfing this plant, 2 novel genotypes carrying either 3 copies of Div^+ or 2 copies of *div* and 1 of Div^+, were identified. Interestingly, the plants with 3 doses of Div^+ had a novel phenotype, a weak ventralisation of the flower, which may reflect over-expression of *Div*.

Selfing Div^+/div plants also gave rise to two other mutants: *centroradialis* and *alien*. The *alien* mutation has not been previously observed in *Antirrhinum* and causes petals and stamens to develop in the regions where ovules would normally occur, whilst the outer part of the ovary, style and stigma as well as the three outer whorls (sepals, petals and stamens) remain similar to wild type. By Northern analysis of dissected flower organs from *alien* and wild type, the Lisbon group has shown that the transcription patterns of the *plena* and *deficiens* genes are similar in the three outer whorls of both genotypes. In the fourth whorl of *alien*, however, *deficiens* transcript levels are greatly increased and *plena* transcripts moderately decreased relative to wild type. *In situ* hybridisation revealed that at a stage when third whorl stamens are visible as small mounds, the patterns of *deficiens* and *plena* expression in *alien* mutants are similar to those of wild type. At a later stage, when the outer

and inner parts of the fourth whorl have become morphologically distinct, *deficiens* expression is very strong on the inner part of the pistil of alien. In the same region of wild type, *deficiens* is not expressed. These results suggest that *alien* identifies a late-acting function which is required for carpel development and acts, at least in part, by negatively regulating *deficiens*.

3. Construction of an RFLP map of *Antirrhinum*

The previous section illustrates how the ability to localise genes on a genetic map can be of enormous value for gene isolation and analysis. Although a good genetic map of *Antirrhinum* exists, it is of limited value for the rapid assignment of map positions to newly cloned DNA sequences. To circumvent this problem, the Köln group has constructed a combined genetic and Restriction Fragment Length Polymorphism (RFLP) map for *Antirrhinum*. Crosses were carried out between inbred lines of *Antirrhinum majus* and *Antirrhinum molle*, two species that are likely to reveal RFLPs. 96 F2 progeny from these crosses have been probed with more than 100 molecular markers (these include randomly selected cDNA clones and clones of known loci, which were provided in part by the Norwich group). This analysis has allowed the map position of 100 cloned DNA sequences to be determined. Because some of these cloned sequences correspond to previously mapped genetic loci, it has also been possible to start aligning the RFLP and classical genetic maps. So far, five out of the eight *Antirrhinum* chromosomes, mapped in the past by classical methods, have been assigned to RFLP linkage groups. The validity of the RFLP map was confirmed by comparison of linkage data obtained by analysis of two independent F2 populations. The map produced has already proved invaluable to all *Antirrhinum* scientists.

4. Construction of double mutants and EMS mutagenesis

The analysis of how genes interact to control flower development can be greatly aided by studying the phenotypes and properties of plants that carry mutations in several genes (e.g. double or triple mutants). The Norwich, Köln and Lisbon groups are involved in this aspect of the programme. For example, the Norwich group has constructed and analysed: *cyc;ovu, def;ple, fim;ple, fim;def, squa;fim, glo;ple, flo;glo, flo;squa, cyc;dich, cyc;rad, rad;cen.* The Lisbon group has made *cyc-;div, cyc;dich, cen;div* and *cyc;dich;div* mutant combinations. The phenotypes obtained have had important implications for models of how flower shape and organisation is controlled. Many other mutant combinations have also been made in the collaborating laboratories. In addition, screens for mutations that suppress or enhance the phenotype of genes such as *flo* (Norwich) or *def* (Köln) have been initiated.

To investigate how the spectrum of mutations obtained from transposon-mutagenesis compares to that from chemical mutagenesis, an EMS-mutagenesis programme was initiated by the Köln group. An EMS-treated M1 generation of 1000 plants was selfed and progeny from about 360 of these were analysed for segregating mutations. 19 mutants affecting flower or inflorescence morphology were obtained. The phenotype of most of these resembled previously described mutants (e.g. *plena, deficiens, obteca, demidensa, divaricata, choripetala, incomposita, unilabiata, eramosa, squamosa).* Three mutant, *lacrima, oclusa* and *desnuda*, display novel phenotypes. Mutants affecting many other characters, such as leaf shape, chlorophyll and embryo development were also obtained and have been distributed to interested parties in Europe. The overall conclusion is that there appears to be extensive overlap in the spectrum of mutants obtained by chemical and transposon mutagenesis.

5. Transformation of *Antirrhinum*

The ability to modify and introduce genes into plants provides a very useful tool for the analysis of gene function. Although genes isolated from *Antirrhinum* can be introduced into related species, such as tobacco, it is not yet possible to re-introduce them into *Antirrhinum*. This is because there is no routine method for transformation and regeneration of *Antirrhinum*. One aim of the programme is to try and develop such a method.

Initial work of the Birmingham group centred on the use of *Agrobacterium tumefaciens* and various experiments were carried out to optimise T-DNA transfer. However, because of the difficulties of using this system with *Antirrhinum*, they switched to using *Agrobacterium rhizogenes*. This bacterium induces the production of hairy roots rather than tumours but, for some plant species, it has been possible to regenerate plants from these transformed roots. A protocol has been developed that allows the production of hairy roots from a wide range of *Antirrhinum* varieties. So far, it has only been possible to induce these roots to form on excised hypocotyl segments. Using *A. rhizogenes* that contains a binary vector, it has been possible to produce GUS-positive hairy root clones which can also be selected on plates or in liquid culture on the basis of their kanamycin resistance. A low rate of shoot regeneration from some hairy root clones has been observed for a few varieties of *Antirrhinum*, the highest rate being from the variety Golden Monarch. A number of hairy root-derived shoots of this variety have been successfully transferred to compost and grown in a glasshouse. They exhibit the classic distorted morphology described previously for other species. The plants are dwarfed because of the very short internodes, the leaves are more ovate and the flowering is markedly delayed compared to wild type. For some species it has been reported that polyamine treatment can result in the production of a more normal phenotype, but this could not be achieved for *Antirrhinum*. When the first GUS-positive plant flowered, it was crossed with wild type. Of 244 progeny, 115 were dwarf and GUS-positive, and 129 showed normal morphology and were GUS-negative. The results indicate a tight linkage between the T-DNAs introduced from the Ri plasmid (bearing the genes responsible for hairy root induction and the dwarfed shoot phenotype) and the disarmed binary plasmid (carrying GUS). The Birmingham group carried out a range of experiments to optimise transformation efficiency, such as using *A. rhizogenes* lines in which one of the *rol* genes has been inactivated or by varying hormone balance in the culture medium. Neither of these approaches greatly improved efficiencies. However, using silver nitrate at concentrations of 10-20 μg/ml did significantly increase shoot regeneration from Golden Monarch hairy roots.

Successful transformation of *Antirrhinum* was also achieved by the Köln group using *Agrobacterium tumefaciens*. From about 400 hypocotyl or leaf explants, two transformed plants were regenerated. The extremely low frequency is due to problems of shoot induction from callus under tissue culture conditions. The transformed plants carry the GUS reporter gene under the control of the CaMV 35 S promoter and a methotrexate resistance marker for selection. The regenerated plants did not show morphological abnormalities and were selfed or crossed to wild type plants. Seeds obtained from one plant after these crosses were germinated under methotrexate selection and showed the expected segregation for the selectable marker. To show the temporal and spatial pattern of the 35S promoter in *Antirrhinum*, the expression pattern of the GUS gene has been examined in one transformed plant.

6. Technology transfer between plant species

Understanding the genetic control of flower development in *Antirrhinum* should provide a model for studies in other species. To investigate the possibility of extending expertise to other plants, the Valencia group is using *Antirrhinum* DNA probes for the analysis of genes controlling flower development in pea. The genetics of pea has been very well researched and many mutants are available. One of these, *deficiens*-like, gives flowers in which carpels (female organs) grow in place of stamens and petals are sepaloid; a phenotype rather like that of the *deficiens* and *globosa* mutants of *Antirrhinum*. In another mutant, *stamina-pistilloida*, only the two marginal stamens of the staminal tube are transformed into pistilloid organs and the petals forming the keel are not fused, showing green sectors and sepaloid shapes in the edges. The Valencia group has completed the morphological characterisation of these mutants with SEM ontogenetic studies. They have also used probes from the MADS-box gene family of *Antirrhinum*, which includes the *deficiens* and *globosa* genes, to isolate 17 cDNA clones from pea, falling into 5 different classes. Two of these have been characterised. pM1 shows high homology with the *globosa* gene and its transcripts are detected only in petals and stamens. No pM1 transcripts are detected in the *deficiens*-like mutant but they appear to be at normal levels in *stamina-pistilloida*. pM2 transcripts are detected in all floral whorls at early stages of development and appear to be at normal levels in *deficiens*-like but reduced in *stamina-pistilloida*. In addition to isolating MADS-box genes, the Valencia group has also isolated a cDNA, with strong homology with type 2 lipoxygenases, that is repressed during late stages of carpel development. Using three different types of immunisation procedure the Valencia group has also isolated a set of monoclonal antibodies specific for different flower parts (sepal, petal, stamen and pistil) and they have been extensively used in both Western blots and *in situ* immunolocalisation assays to identify tissue mixtures among the original organ and the mutated one. One of the antibodies identifies a protein with high homology with the ALB 2 protein, which is abundant in pea cotyledons. The experience of the Valencia group with immunolocalisation in flowers has also helped their work to be extended to yet another species, *Petunia*. In collaboration with Dr A van Tunen (CPRO-DLO, Wageningen), they have been able to immunolocalise the FBP1 protein of *Petunia*, which has strong homology with *globosa* from *Antirrhinum*. FBP1 appears to be located in nuclei of petal cells and young stamen primordia.

MAJOR SCIENTIFIC BREAKTHROUGHS

(1) Saturation transposon-trapping
(2) Construction of RFLP map
(3) Characterisation of novel gene interactions controlling flower development
(4) Transformation and regeneration of *Antirrhinum*
(5) Initiation of molecular studies on pea flower development.

MAJOR COOPERATIVE LINKS

The success of this programme has depended on the effective combination of complementary skills and resources in the participating laboratories. The disciplines covered include: molecular biology, classical genetics, tissue culture, developmental and comparative biology. Each laboratory alone would have been unable to develop all the tools necessary but in combination they have firmly established a powerful and generally useful model system. The particular exchanges that have occurred are:

1. Meetings

We have had three major meetings, one held in Köln from 17/4/91-21/4/91, one in Valencia from 13/5/92-16/5/92 and one in Norwich from 26/5/93-30/5/93. These involved presentations from all of the participants and have proved to be very useful for exchanges of ideas and results. A report on the first meeting was published in Flowering Newsletter (see below).

2. Materials

Many clones of genes have been freely exchanged between participants, including clones of *flo, ple, def, glo, squa, del, pal, olive, inc.* In addition, cloned transposons exchanged include Tam1, Tam2, Tam3, Tam4, Tam5, Tam6, Tam7, Tam8 and Tam9. Many genetic stocks have also been freely exchanged, including those carrying mutations at *flo, ple, glo, squa, inc, div, pal, def.*

3. Staff exchanges

J. Almeida spent July 1991 working with the Norwich Group. S. Hantke from the Norwich group worked for 3 weeks with the Köln group in December, 1991. M. Mooney and P. McSteen from the Norwich Group visited the Birmingham group for 2 days in September, 1991. E. Andrade from the Lisbon group has joined the Köln group to work on a PhD. P. McSteen and J. Roberts from the Norwich Group visited the Birmingham group in October 1992. M. Rocheta from the Lisbon Group came to work with the Norwich Group for 1 month in March 1993. M. Rodriguez from the Valencia Group spent two months with the Köln Group in 1992. J. Almeida visited the Köln group in September 1993. R. Carpenter and E. Coen visited the Köln group for two days in February, 1994. C. Perbal from Köln visited the Norwich group in March 1994 to discuss *in situ* hybridisation methods.

PUBLICATIONS

Joint publications

Carpenter, R., Coen, E.S., Huijser, P., Schwarz-Sommer, Z., Sommer, H. (1991) The genetic control of flower development in *Antirrhinum*. Flowering Newsletter **12**, 38-41.

Bradley, D., Carpenter, R., Sommer, H., Hartley, N. & Coen, E. (1993). Complementary Floral Homeotic Phenotypes Result From Opposite Orientations of A Transposon at the *plena* Locus of *Antirrhinum*. Cell, **72**, 85-95.

In January 1992, we submitted a report on the programme, published as a brochure by the Commission.

Individual publications

Holford, P., Newbury, H.J. (1992) The effects of antibiotics and their breakdown products on the *in vitro* growth of *Antirrhinum majus*. Plant Cell Reports, **11**, 93-96.

Holford, P., Hernandez, N, Newbury, H.J (1992) Factors influencing the efficiency of T-DNA transfer during the co-cultivation of *Antirrhinum majus* with *Agrobacterium tumefaciens*. Plant Cell Reports, **11**, 196-199.

Simon, R., Carpenter, R., Doyle, S. and Coen, E. (1994). Fimbriata controls flower development by mediating between meristem and organ identity genes. Cell, (in press).

A preface to the final reports of Bridge contracts BIOT CT-900174 and BIOT CT-900172

INTRODUCTION

Self-incompatibility (SI) is an important tool in the breeding of a small number of crops, for it permits the production of F1 hybrid seed without the need for artificial pollination systems or the use of male sterile lines. The prospect of extending the use of SI to other species is commercially extremely attractive, but research leading up to the BRIDGE programme had already identified at least three types of SI system, all regulated by highly-complex S(incompatibility)-loci. The chances of transferring any of these loci to new species were therefore very slight — especially as very little was known of the molecular basis of the recognition and rejection systems involved. Further, even in plants where SI is currently in use by breeders (e.g. *Brassica oleracea*) problems of reliability were being encountered in the field performance of certain alleles. The plant breeding industry was thus of the opinion that the market in the late 1990s would require SI lines which were far more reliable, and available for transfer to new crops. If these commercial targets could not be met, it was clear that industrial interest would be transferred to new natural and engineered male sterility systems.

In response to this perceived requirement for basic research on the molecular basis of SI in a range of key crop plant species, two separate projects were submitted for funding under BRIDGE. One focused on the complex sporophytic SI (SSI) system of *Brassica oleracea*, while the other was centred on the RNAse-based, gametophytic SI (GSI) system which characterises members of the Solanaceae. While each project was multidisciplinary and involved participants from a wide range of member states, it was obvious to the coordinators of both projects (Prof H. Dickinson [Oxford, UK] and Dr Richard Thompson [MPI; Köln] respectively) and also officials in DG XII that — were the projects to be funded — there should be strong collaborative links between them. For this reason the Technical Annexes of each proposal contained sections describing mechanisms by which this collaboration should take place. In brief, these involved identification of key areas of common scientific and technical interest — for example the organisation of S-locus promoters in SSI and GSI systems which seemingly share a number of common features.

As the work progressed collaborative links of a more practical nature were set up, with the production of SINEWS — a newsletter serving the interests of both projects, and the holding of three very successful joint workshops.

Collaboration between the two projects continued at almost every level throughout the duration of the BRIDGE programme, and thus much of the resulting 'added value' can be found in the reports of the individual projects which follow this preface. However, several successful pieces of work arose directly from this inter-project collaboration, and two examples of these are described in the following paragraphs. Details of the joint newsletter and workshops, are also provided (adapted from the main project reports).

INTER-PROJECT RESEARCH

A. Transcription factors regulating the activation of S(incompatibility) promoters.

Work by the group in Birmingham on *Brassica* S-locus related glycoprotein gene promoters had discovered domains likely to be involved in the activation of these sequences in the stigma. To identify the transcription factors responsible, Dr R. Hackett travelled to Dr R. Thompson's laboratory in Köln which was actively developing strategies for isolating such factors in members of the Solanaceae. Supported by a short-term EMBO fellowship Dr Hackett established a method for nuclear extract preparation from styles, based on a standard protocol for preparing nuclear proteins, but modified to deal with the relatively small quantities of stylar extracts and the low concentrations of nuclear proteins present. On the basis of a deletion series analysis of the *Brassica* SLR-1 promoter expressed in tobacco, a 300-basepair promoter region conferring pistil-specific expression had been previously identified and a series of smaller fragments from this region were isolated via PCR, end-labelled and used to screen nuclear extracts for sequence-specific interactions using the electrophoretic mobility shift assay (EMSA) (see Figure 1). Three such complexes were detected, demonstrated to be sequence-specific, and the binding sites involved were further narrowed down using smaller fragments as probes. Since returning to Birmingham, Dr Hackett has further improved the quality of the nuclear extracts, and narrowed down protected sites suitable for use in screening an expression library. In Köln, a technique for screening for DNA-binding-factor clones in yeast has been established, and is being applied to motifs derived from the pistil-specific promoter fragments.

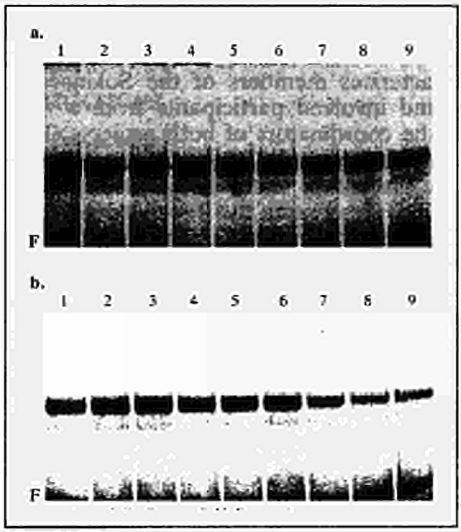


Fig. 1
Gel shift assays using tobacco pistil cell extract in binding reactions with a labelled 48mer from within the pistil-specific region of the Brassica SLR1 gene promoter.
a. Lane 1: interaction of a binding activity in pistil cell extract with the labelled 48mer in the presence of 2µg of the synthetic co-polymer p(dA-dT).p(dA-dT). Lanes 2 to 9: as lane 1 with the addition of increasing molar excess of a 51mer from elsewhere in the SLR1 promoter, as a non-specific competitor. Molar excess ranges from 1-fold (lane 2) to 150-fold (lane 9).
b. Lane 1: as lane 1 gel 'a'. Lanes 2 to 9: as 'a' with the addition of increasing molar excess of unlabelled 48mer as a specific competitor. Molar excess ranges from 1-fold (lane 2) to 150-fold (lane 9).
F: position of free probe.

B. Identification of putative S-related sequences in sugarbeet

The **Maribo Seeds** group are studying the molecular basis of SI in sugarbeet. Dr Petersen has undertaken various approaches to search for a candidate molecular marker, including protein gel electrophoresis of style extracts from SI and SC beet lines, construction of a style cDNA bank and isolation of pistil-specific DNA clones, isolation of PCR clones using oligonucleotide primers based on sequences conserved either in the SLGs of *Brassica* or in the S-RNases of the Solanaceae. In the course

of this work, Dr Petersen received a number of samples and protocols from Köln, and worked for short periods in the Köln lab. Currently, the most promising data have been obtained from three cDNA clones related to SLGs isolated from the style cDNA bank. One of these corresponds to an SRK (the *Brassica* S[incompatibility]-specific receptor kinase) in sequence, and is a candidate for an S-locus component, although genetic linkage has still to be carried out — a suitable segregating population not yet being available.

'SINEWS' — THE INTER-PROJECT NEWSLETTER

An important outcome of the first workshop in Enkhuizen was the decision to produce SINEWS — a newsletter to which participants from both projects contributed. SINEWS was published every 6 months and contained research reports from all participants, registers of expertise and materials available for exchange (see Final report of BIOT CT90-0174[SMA] below), and lists of training opportunities available within each project. SINEWS also carried accounts of scientific meetings of general interest and contained the reports made to the Commission at the end of each financial period. So far there have been six issues of SINEWS, one further edition will be produced later in 1994. Back copies of SINEWS are available from the coordinating laboratory of the *Brassica* SI project (Oxford).

INTER-PROJECT WORKSHOPS

There have been three BRIDGE workshops over the course of the programme, the first of which was held in April 1991 at Enkhuizen, hosted by Zaadunie BV (now S&G Seeds). At this initial meeting all participants presented their proposed research strategies and decisions were made as to the shape and future coordination of the two projects. Two further collaborative meetings have since taken place; in April 1992 both groups met in Köln at a workshop hosted by the MPI and organised by Dr R Thompson. Dr T-H Kao (Pennsylvania State University) gave the plenary lecture, thus placing the EC-supported work in a fully international perspective. The last workshop was hosted by the Ecole Normale Superieur in Lyon and organised by Dr T Gaude. The plenary lecturer on this occasion was Dr J Walker (University of Missouri) who spoke on transmembrane kinases — a subject of current interest to participants from both projects. Members of the Plant industrial Platform were invited to attend the second day of this meeting and heard the plenary address, a summing up of the two projects (presented by Dr R Thompson and Prof H Dickinson) and a presentation on behalf of the Commission by Dr E Mangien.

Throughout the duration of the programme, small *ad hoc* meetings have been organised for specific purposes (e.g. to coordinate transformation strategies).

THE 'ADDED VALUE' OF COLLABORATION BETWEEN PROJECTS

A list of publications, or even 'research breakthroughs', does not really provide an accurate picture of the achievements of a large transnational research programme. Key to the real success of the work, and indeed of research in ensuing years, are the links that are forged during an extended period of collaborative research. The conjunction of the two projects, with different foci but many aims in common permitted the formation and development of many such links; certainly some had formed between the participants of individual projects prior to BRIDGE, but the active collaboration fostered by the workshops, SINEWS and other mechanisms contributed very significantly to the 'added value' of this highly-successful EC sponsored research programme.

The molecular basis of cell-cell interactions in self-incompatibility (BIOT CT-910172)

COORDINATOR:
R.D. THOMPSON, MPI für Züchtungsforschung, Köln, D
PARTICIPANTS:
G.J. WULLEMS, University of Nijmegen, Toernooiveld, Nijmegen, NL
M. CRESTI, Universita degli studi di Siena, Siena, I
M. van GRINSVEN, S&G Seeds, Plant Biotechnology Group, Enkhuizen, NL
G. WRICKE, Universität Hannover, Hannover, D
R. THOMPSON, MPI für Züchtungsforschung, Köln, D
E. JACOBSEN, Univ. Wageningen, Wageningen, NL

BACKGROUND INFORMATION

Self-incompatibility (SI) is a naturally occurring phenomenon giving rise to F1 hybrid seed, which is potentially widely applicable in plant breeding. Currently, the S-locus, which controls SI, is incompletely characterized at the molecular level in any plant species. It is however already evident that a number of fundamentally distinct SI mechanisms exist, the sporophytic SI operating in *Brassica* involving different components to the gametophytic SI of solanaceous species, for example. To date, the putative S-gene products expressed in the stylar tissue have been identified, but there is an urgent need to identify the corresponding S-gene products expressed in pollen, and to elucidate the mechanism of interaction between these components which results in self-sterility. The results will facilitate the application of SI in plant breeding, and also help us to define a unique cell-cell recognition event in plants.

OBJECTIVES AND PRIMARY APPROACHES

The project was aimed at improving our basic knowledge about potato SI and and at providing materials for a complete molecular characterization of a gametophytic S-locus. Further, a comparison was to be made to the two-locus gametophytic SI system of Rye. Primary approaches were the characterization of genomic and cDNA clones for the S-locus, a comparison of different S-alleles, and the effect of a self-compatible mutant on expression of the S-gene products. Further groundwork was to be provided by the development of vigorous lines of known S-genotype, and lines amenable to transformation. A multi-headed strategy to the isolation of possible pollen-S components was taken, with the screening of pollination-induced cDNAs from pollinated pistils, the isolation of cDNAs preferentially expressed in germinated pollen, and latterly, a map-based cloning strategy based on screening tomato YACs flanking the S-locus for pollen-expressed sequences.

RESULTS AND DISCUSSION

In **Nijmegen**, two pistil-specific genes, STS14 and STS15, were isolated from cross-pollinated pistils of Solanum tuberosum. The expression of STS15 is predominantly in the stylar cortex as shown by *in situ* hybridization analysis. Other isolated genes expressed in pistils are involved in the flavonoid biosynthetic pathway. Among those are the genes encoding flavonol synthase (FLS) and isoflavone reductase (IFR). The accumulation of the transcript encoding isoflavone reductase increases after pollination in pistils of potato.

Transcripts of the pollen-specific gene ntp303 accumulate during the late phase of pollen development and during pollen tube growth in tobacco pistils. Northern

analysis in pollen from mono- and di-cotyledonous plant families showed that this gene is evolutionarily conserved in different families. Furthermore, by gain-of-function and deletion analysis, we demonstrated that the sequence AAATGA of the ntp303 regulatory region is necessary and sufficient for the pollen specificity of gene expression.

The functional analysis of the gene ntp303 and the isolated flavonoid genes (IFR and FLS) may help to understand the mechanism of pollen tube growth. The identification of this process will provide new data on fertilization and in general on intercellular signalling processes in plants.

In **Siena,** the distribution of the S-RNases S1 and S2 has been established in potatoes of known S-genotype using antibody prepared from *E. coli*-expressed glutathione S-transferase fusions to S1 and S2 coding sequences. S1 and S2 were localized at the surface of the secretory cells in the stylar transmitting tract in the intercellular matrix and in the papillar cells of the stigma. The pistil-specific endochitinase SK2 was localized using antibody prepared using FPLC-purified protein. As neither SK2 nor the *E. coli*-expressed S1 and S2-RNases were glycosylated, a polyclonal antibody could be used for immunochemical localization without background problems. However, for future purposes and for reference, monoclonal antibodies were also raised to SK2 and characterized with the antigen. SK2 was localized by immunofluorescence microscopy in the transmitting tract and in the cytoplasm of stigma secretory cells underlying the papillae. Transgenic plants expressing S2-RNase in pollen were developed to test whether the gene product in pollen is modified in some way that would distinguish it from the stylar S-RNase. The protein accumulated in pollen although no biochemical modification could be detected. However, at the ultrastructural level, the protein was seen to be localized primarily in the cytoplasm, particularly close to the plasma membrane, in contrast to its extracellular location in the style, suggesting an altered targeting event may have occurred in the transgenic plants.

In **Enkhuizen** a tomato transformation protocol was developed which employs cotyledons from young *in vitro* cultured seedlings. Prior to the actual co-cultivation with the *Agrobacterium* suspension a preculture of one day of the explants on a tobacco feeder-layer is necessary. During the 2 day co-cultivation a tobacco feederlayer is used to improve the transformation frequency. Using this protocol we typically observe shoot formation on 15% of the explants. Unlike reports from other groups on transformation of tomato or other crops, all these selected shoots are true transgenics. nevertheless, we do observe that only approximately 50% of the thus-selected shoots have a diploid nuclear structure. The reminder is either chimeric, tetraploid or even octoploid. This leads in total to a general transformation efficiency of approximately 5% of the originally co-cultivated explants. Seed set on these transgenic plants is normal. Collection of seeds is typically between 7-9 months after initiation of the experiment.

At this moment the transformation of tomato lines is rather straightforward and reproducible. It has been tested on several elite-genotypes of tomato lines, representing the complete range from fresh market tomatoes to industrial bush tomatoes. Although our protocol is widely applicable, the overall transformation frequency can vary from genotype to genotype. A striking observation is the fact that different seed-batches from the same genotype can lead to different overall frequencies. We do not have a good explanation for this observation.

With respect to the molecular structure of the inserted genes it is clear that next to single insertion events, multiple insertion events do occur frequently amongst our original transformants. In following generations a part of these multiple insertions behaves as unlinked loci, whereas the reminder appeared to be closely linked. It is expected that these latter cases represent tandem insertions. The majority of the selected transformants do transmit their 'transgene' faithfully to the next generation. Nevertheless, a few lines have been identified that appear to have lost expression of the transgene. Whether this is due to methylation of the inserted sequence or other 'epigenetic' effects is not investigated.

At this moment we are evaluating a number of tomato transformants of the S3 generations and experimental hybrids in field trials. We conclude that the tomato-transformation protocol developed at S&G Seeds is suitable for the purpose initially out lined for this BRIDGE program.

The results obtained in **Hannover** include:

a) ***SI-related proteins in Rye.*** Rye pistil proteins from defined S-genotypes were fractionated by IEF and 2D gel-electrophoresis and inspected for the presence of S-linked polypeptides and S-linked RNase species, but none were detected. Similarly, amplification by PCR of S-RNase type sequences from stylar RNA was unsuccessful. The SI-reaction in Rye pollen is characterized by a rapid protein phosphorylation response. A major phosphate-labelled protein differed after *in vitro* treatment of pollen with either 'cross' or 'self' stigma extracts, with strong phosphorylation occurring only with 'self' extracts. Self-compatible (SC) mutants lost this response, and the reaction was also abolished by application of either protein kinase inhibitors such as Lavendustin or the Ca^{2+} antagonists Verapamil and La^{3+}.

b) ***S-locus linked DNA fragments.*** Rye inbred lines segregating at the S-locus but homozygous at the Z-locus were used as templates for PCR with primers based on *Brassica* SLG sequences. A 280 bp fragment possessed a polymorphism linked to S on DGGE fractionation of segregating populations totalling 46 plants.

The results obtained in **Köln** include:

a) ***S-RNase gene isolation:*** Stylar proteins from a series of dihaploid lines classified for their S-genotype by Prof. Hermsen (Wageningen) were fractionated by 2D-SDS-IEF electrophoresis and used to associate one basic glycoprotein to each of the S-alleles S1 to S4. Two further abundant proteins, SK1 and SK2 were present irrespective of S-allele constitution. Using primers based on protein sequence for S1-S3 and SK2, cDNA and genomic clones have been obtained and sequenced. The cDNA clones for S1 and S2 have been expressed as GST-fusions in *E. coli* and used to raise antibodies for immunocytochemistry. The putative promoter regions of S1 and S2 and SK2 have been tested for their ability to confer style-specific expresion by particle gun bombardment of styles and in transgenic tobacco plants. Attempts to express the entire S1 gene in transgenic potato and tobacco failed; the transformants had no detectable transcript. Reporter genes using the S2 promoter fused to P-glucuronidase, neomycin phosphotransferase or chloramphenicol acetyltransferase (GUS, NPTII, CAT) displayed similar expression patterns in transgenic plants: enzyme activity in the pistil, either in the transmitting tract, at the stigma or in both locations, and a much higher (> 10-fold specific activity) enzyme accumulation in later pollen development, although no S2 transcript could be detected by PCR in potato pollen. A deletion analysis of the S2 promoter showed sequences

conferring pollen and stylar expression are located within the first 300 bp upstream from the transcript start. This region is now being more intensively dissected to identify sequence motifs responsible for this expression.

b) ***Attempts to identify S-sequences expressed in the male gametophyte.*** The identity of the female-expressed S-component as the S-RNase has now been confirmed by transgenic plant experiments. Genetic analyses of gametophytic SI suggests that the male-expressed sequence cosegregates with the female-expressed sequence. Therefore we have used PCR to search for S-RNase transcripts in developing anthers and mature pollen, but have not detected such transcripts in potato. Assuming the male S-locus component is expressed from a closely-linked second open reading frame, the chromosomal region around the S-locus has been isolated in YACs from the closely related species *L. esculentum* (cultivated tomato). As this species is self-compatible, and the male transcript might not be expressed, an SI wild relative accession *(L. hirsutum)* has been identified which possesses a closely related S-allele. It is now planned to screen a cDNA bank from pollen of this species with YAC DNA from the L. esculentum S-locus.

In **Wageningen**, SI and SC clones including all homozygotes and heterozygous conbinations of the S-alleles S1 through S4 were generated, except for the SC S1 homozygote, which may not be possible due to an allele-specific effect of the self-compatibilizing factor. From these clones, which are maintained as tubers, large seed lots were produced by intercrossing and harvested for long-term storage.

A series of readily transformable and regenerable genotypes have been selected out of crosses with efficiently transformable SC genotypes possessing S3 and Sx, an unidentified S-allele. Backcrosses have been carried out to segregate Sx and the compatibilizing factor from the lines. The best line produced has an estimated transformation efficiency of 90% of the regenerated calli after 8 weeks on kanamycin selection. The analysis of potato plants transformed with S2-RNase constructs is still in progress but preliminary results suggest that the expression of the S2-RNase-containing transgenes is too low to give a phenotype. Three S2-promoter/reporter constructs give similar expression patterns, with the highest activity being seen in pollen, and much lower and variable levels in the pistil, but also in other floral tissues and non-floral meristematic tissues, i.e., broadly similar results to those seen in transgenic tobacco.

MAJOR SCIENTIFIC BREAKTHROUGHS

The identification of a tightly linked DNA marker for S in Rye which should facilitate its application in hybrid breeding, and is a useful tool to characterize the locus molecularly.

The identification of efficiently transformable SI genotypes of known S-allele composition provides an indispensable tool for potato SI research.

MAJOR COOPERATIVE LINKS

The accumulation of transcripts growing pollen tubes could be confirmed.

A Bartalesi and X.Q. Li both worked in Köln for short periods to isolate protein and fix material from defined potato lines. Antibodies and transgenic plants were provided to Siena by the Köln group.

PUBLICATIONS

Joint publications

Kaufmann, H., Kirch, H.H., Wemmer, T., Peil, A., Lottspeich, F., Uhrig, H., Salamini, F., Thompson, R. (1992) Sporophytic and gametophytic self-incompatibility. In: Sexual Plant Reproduction (M. Cresti, A. Tiezzi, eds.) Springer-Verlag, ISBN 3-540-55746-6, pp. 115-125.
Kaufmann, H., Salamini, F., Thompson, R.D. (1991) Sequence variability and gene structure at the self-incompatibility locus of *Solanum tuberosum*. Mol. Gen. Genet. **226**: 457-466.
Y.Q. Li, C. Faleri, R.D.Thompson, A.Tiezzi, R.Eijlander, M. Cresti (1994) Cytochemical localization of the abundant pistil protein SK2 in potato (*Solanum tuberosum*). Sex. Plant Reprod. 7, 164-168.
Li, Y.Q., Faleri, C., Thompson, R.D. Tiezzi, A., Eijlander, R., Cresti, M. (1994) Cytochemical localization of the abundant pistil protein SK2 in potato (*Solanum tuberosum*). Sex. Plant Reprod. 7, 164-168.

Individual publications

G.J. van Eldik, W.H. Vriezen, M. Wingens, R.K. Ruiter, M.M.A. van Herpen, J.A.M. Schrauwen, G.J. Wullems. A pistil-specific gene of *Solanum tuberosum* is predominantly expressed in the stylar cortex. (submitted to Sex. Plant Reprod.)
K.A.P. Weterings, W.H. Reijnen, G. Wijn, K.J.P.T. van den Heuvel, N. Appeldoorn, G. de Kort, M.M.A. van Herpen, J.A.M. Schrauwen, G.J. Wullems (Sex. Plant Reprod. 1994,in press) Molecular characterization of the pollen genomic clone ntp303 and *in situ* localization of expression.
Wehling, P., Hackauf, B., Wricke, G., (1994) Phosphorylation of pollen proteins in relation to self-incompatibility in Rye *(Secale cereale L.)* Sex Plant Reprod. 7: 67-75
Wehling, P., Hackauf, B., Wricke, G., (1994) Characterization of the two-factor self-incompatibility system in *Secale cereale* L. in Hybrid Breeding eds., G. Wricke and U. Kück, pub. Blackwell, Edinburgh.
Wehling, P., Hackauf, B., Wricke, G., (1994) Identification of S-locus linked PCR fragments in rye (*Secale cereale* L.) by denaturing gradient gel electrophoresis. Plant J. in press.
Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R.D., Bonierbale, M.W., Ganal, M.W., Tanksley, S.D., Salamini, F. (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor. Appl. Genet. **83**: 49-57.
Thompson, R.D., Kirch, H.H. (1992a) The S-locus of flowering plants: when self-rejection is self-interest. Trends in Genetics **8**: 381-387.
Thompson, R.D., Uhrig, H., Hermsen, J.G.Th., Salamini, F., Kaufmann, H. (1991) Investigation of a self-compatible mutation in *Solanum tuberosum* clones inhibiting S-allele activity in pollen differentially. Mol. Gen. Genet. **226**: 283-288.
Wemmer, T., Kaufmann, H., Kirch, H.H., Schneider, K., Lottspeich, F., Thompson, R.D. (1994) The most abundant soluble basic protein of the stylar transmitting tract in potato (*Solanum tuberosum L.*) is an endochitinase. Planta (1994) **194**: 264-273
Colon, L.T., Eijlander, R., Budding, D.J., Ijzendoorn, M.T. van, Pieters, M.M.J., Hoogendoorn, J. (1993) Resistance to potato late blight (*Phytophthora infestans* (Mont. de Bary) in *Solanum nigrum, S. villosum* and their sexual hybrids with *S. tuberosum* and *S. demissum.* Euphytica **66**: 55-64.
Eijlander, R., Stiekema, W.J. (1994) Biological containment of potato (*Solanum tuberosum*): outcrossing to the related wild species black nightshade (*Solanum nigrum*) and bittersweet *Solanum dulcamara*). Sex. Plant Reprod. 7: 29-40.

Molecular genetics and physiology of self-incompatibility in *Brassica* crops (BIOT CT-900174)

COORDINATOR:
H.G. DICKINSON, Oxford University, Oxford, GB
PARTICIPANTS:
C. DUMAS, INRA, Lyon, F
R. CROY, University of Durham, Durham, GB
G. WULLEMS, University of Nijmegen, Nijmegen, NL
E. OOST, Nickerson-Zwaan B.V., Barendrecht, NL
M. VAN GRINSVEN, Zaadunie (S&G) B.V. , Enkhuizen, NL
F. OKKELS, Maribo Seeds, Copenhagen, DK
M. TRICK, JICPSR, Cambridge Lab., Norwich, GB
C. FRANKLIN, University of Birmingham, Birmingham, GB
D. OCKENDON, HRI, Wellesbourne, GB
L. DAVIDSE, Royal Sluis B.V., Enkhuizen, NL

BACKGROUND INFORMATION

In discussion with industry, a number of serious bottlenecks impeding progress towards the successful commercial application of SI in *Brassica* breeding were identified. Chief amongst these were the unreliability of S-alleles under field conditions and the need to transfer S-alleles between plant lines. Clearly, these problems could only be resolved once SI was explicable in cell and molecular terms, and for this reason a highly-focused coordinated research programme was devised, involving four key industrial concerns and the best six European laboratories in the field. The programme submitted to the Commission aimed to exploit the multidisciplinary skills available in the participating laboratories as well as the important 'commercial dimension' provided through the links with industry. In a parallel application (BIOT-90-0172), Dr R. Thompson and collaborators were proposing to study SI in solanaceous crops and the Commission requested that the activities of the two groups should be coordinated. This was carried out through joint workshops, a newsletter and active collaboration in infrastructure activities, such as establishing databases and raising antibodies. **An informal joint report will shortly be published, highlighting the progress made in these collaborative endeavours.**

OBJECTIVES AND PRIMARY APPROACHES

To exploit the multidisciplinary skills available in the participating laboratories and to promote efficient management, the programme was subdivided into 4 'Target Themes' and a contracting laboratory made responsible for managing each theme. The general area covered by each theme, the contractors managing them and the potential 'deliverables' resulting from the research are listed below:

1. Male component of S-locus (Manager: Lyon).
   In vivo bioassay for SI in *Brassica*; identity of the male determinant of SI; Organisation of S-locus in vicinity of male-expressed gene(s).

2. Operation of self-incompatibility (Manager: Oxford).
   Cell-biological characterisation of the pollen/stigma interaction; 3D analysis of the interactions between male and female molecules; physiology of the rejec-

tion/acceptance response; role of pollen components in SI; physiology of S-allele variability.

3. Structure of S and S-related loci (Manager: Norwich).
   Promoter analysis of S-locus genes; identification of the domains of S-locus genes conferring S-specificity; an understanding of the roles of S-locus linked and related glycoproteins (SLGs and SLRs).

4. Relationship between S-alleles (Manager: Norwich).
   The molecular basis of allelic 'strength' and 'weakness'; molecular characterisation of a range of S-alleles; S-allele diagnostic probes.

5. Transfer of S-alleles (Manager: Norwich).
   Reliable transformation of *Brassica oleracea*; transfer of working SI systems to new lines: evaluation of transgenic lines as potential crops.

RESULTS AND DISCUSSION

Excellent progress has been made under nearly all of the theme headings with most of the objectives being achieved (see below). The only exception was Theme 4, where extreme difficulties were experienced with transforming *Brassica oleracea* (See under 'Major cooperative links').

Identification and characterisation of the male component of the S-locus

Work at Oxford has produced both an *in vivo* bioassay for the male determinant of SI, and also a promising candidate for the determinant itself. Using a combination of HPLC, FPLC and micromanipulation it has very recently proved possible to determine which components of the pollen grain coating play a part in SI. The method has already demonstrated the determinant to have a low Mr and to be carried within the coating. These data accord well with the previous discovery of a 7kD coating borne peptide which binds specifically to female-expressed S-locus products. These discoveries have been exploited in collaborative work with Durham in which N-terminal sequence data from the from pollen coating peptides (PCP) have enabled the isolation of a cDNA clone (PCP7) using PCP-specific primers in a 3' RACE PCR technique. This clone has provided the remaining amino acid sequence of the protein and a probe for the isolation of the genes. *PCP* has been shown to be expressed as a 400 nucleotide mRNA; expression is anther specific and only in the late stages of anther development (between 7mm and 10mm bud length), ceasing after flower opening. *In situ* hybridization experiments are under way to discover the precise localization of the expression of *PCP* genes. The relationship between *PCP* and genes of the S-locus is discussed below.

To determine whether the SLG mRNA expressed in male tissues differs from that expressed in the pistil, the group at Lyon amplified by PCR and cloned SLG cDNAs from anther mRNA. DNA sequencing revealed no differences between anther and pistil SLG cDNAs. A procedure involving extraction of mRNA from single organs was developed in order to rule out the possibility that the SLG mRNA detected in anthers was due to a low level of contamination with pistil tissue. RACE-PCR has been used to analyse the expression of SRK in male tissues. A new expressed member of the S-gene family has been identified: this gene is unusual in that it contains homology with both the kinase and S-domains of SRK but is predicted to encode a secreted glycoprotein similar to SLR2. This gene has been designated SLR3. The identification of the SLR3 gene has been achieved in collaboration with Dr Borislav Stanchev (Durham).

Early attempts at Durham to produce *Brassica oleracea* large fragment libraries using YAC and P1 cloning vectors were only partially successful and did not produce sufficiently large libraries to enable isolation of S-locus clones. The group has, however, successfully screened a conventional S29 lambda library constructed by Vinod Kumar in Martin Trick's group (Norwich), using PCP, SLG and SRK probes. This has yielded about 30-40 PCP-positive clones as well as a number of SLG/SRK positive clones. To date two of the *PCP* genomic clones have been sequenced entirely, including regions about 2-3 kb upstream of the known coding sequences. There are only minor differences between the gene and cDNA coding sequences. Surprisingly in both cases the gene sequence data show no obvious initiation codon in the immediate proximity of the known coding region, thus suggesting the presence of at least one or perhaps two introns. 5'- RACE experiments are in progress now in order to find the transcription start of the *PCP* genes.

Using methods developed in Oxford to isolate pollen coatings, the group in Nijmegen have cloned genes encoding a number of polypeptides carried in the pollen coating. Sequence analysis of the cDNAs obtained shows a number of important homologies. One corresponds to a gene encoding profilin, an allergenic actin-sequestering protein while another has strong homology with oleosin. Profilin and oleosin-like proteins seem to accumulate predominantly in the pollen coat. Profilin may also be involved in the intercellular signalling system between pollen and stigma, while the oleosin-like protein may have a function in the hydration process preceding pollen germination. A third homology was with water-stress modulated genes; interestingly none of these genes was induced in leaves upon stress — suggesting that these sequences may be pollen specific. F_2 segregation analysis showed that none of the genes corresponding to these cDNAs were linked to the S-locus.

Biochemistry and physiology of SI response

In an investigation into the biochemical mechanism of the SI response the Durham group used okadaic acid, a specific inhibitor of serine/threonine protein phosphatases (type PP1 and PP2A) to demonstrate their involvement in SI. Detached pistils from *B. oleracea* S29 and S63 inbred plants were incubated in okadaic acid for defined periods and then tested for self-pollination. Non-toxic concentrations of the inhibitor caused the breakdown of SI in both S63 and S29 lines providing evidence for the involvement of protein phosphatases and therefore the likely involvement of signal transduction mechanism in SI and pollination. A direct connection between the S-receptor protein kinase (SRK) and the protein phosphatases revealed by these experiments has not been established but is worthy of speculation. This work has been published. In order to study and manipulate SI in the field, research at Nickerson has focused on

(1) developing a test method to discriminate between *Brassica oleracea* lines of differing 'strengths' of SI and
(2) finding a method that modifies the expression of strong SI in such a way that seeds from inbred lines can be produced efficiently.

The incompatibility score (IS) system was devised to evaluate allelic strength and proved very valuable when studying both absolute allelic strength and interactions with the environment. To obtain self seed from strongly SI lines, methods investigated included both salt and heat treatments; both worked well and offer very effective alternatives to the current approaches of CO_2 treatment and bud pollination.

Cell biological studies at Oxford have revealed that the SI response is highly focused on the contact point between the self-grain and the stigma surface, and also that the stigmatic papilla is capable of allowing the developing of a cross-grain whilst rejecting a self. Detailed studies of the signals emanating from the pollen coat and the response at the stigmatic surface have identified three distinct interactions, of which only one is regulated by the S-locus. Importantly, all three pathways are also activated within somatic epidermal cells when challenged by fungal pathogens. *Brassica* SI is thus emerging as a useful experimental paradigm for the far more complex host/pathogen interactions.

At Lyon a thorough analysis of the expression of SLG glycoproteins in pistil tissues of the self-compatible P57Sc line has been carried out and has demonstrated that SLG-*Sc* glycoproteins are expressed, although at a low level, in the transmitting tissue and ovary. Their function remains to be investigated.

Organisation and expression of the S-locus, relationship beteween S-alleles

The relationship between male and female expressed S-locus genes

In Durham the screening of the S29 genomic library with PCP (male peptide) and SRK (female S-linked receptor-kinase) specific genes has revealed one extremely interesting clone which contains both a *PCP* gene and an SRK-like gene. This close physical linkage between two genes implicated in SI through functional properties and sequence homology is highly suggestive and, while this falls short of definite proof of identity, is worthy of further investigation with a large-fragment genomic library. Such a cosmid library is currently in preparation for screening. The PCP-linked SRK gene has an interesting and unique structure.

Also at Durham, major effort has been expended on trying to establish any S-locus linkage of the *PCP* genes through the use of RFLPs and F2 segregating populations of progeny from two different S-lines. The initial Southern blot experiments carried out with only two genotypes, S5 and S29, failed to show any polymorphic patterns using PCP7 as a probe and several different restriction enzymes. In collaboration with the Wellesbourne group (see separate report) we have been able to extend our search using a large range of different S-lines. We have as a result been able to reveal at least one definite PCP7 — RFLP and experiments are now underway to analyse progeny from S-lines exhibiting the different patterns. Reconstruction experiments have shown that the copy number of PCP7-related genes varies significantly between different lines, from near single copy up to as many as 50 copies per haploid genome. The significance of this variation in connection with SI and the S-locus is as yet unclear. *PCP* genes appear to be restricted only to the *Brassicas*.

The SRK gene family

At Norwich, SLG and SLR1 cDNAs had been previously cloned from the S29 genotype of *Brassica oleracea*. The SLG29 probe was used to screen an S63 genomic library and a positive clone was isolated. This was found to contain a receptor kinase gene similar to the S-linked SRK2 and SRK6 sequences, enabling further studies of the SRK gene family. A pair of PCR primers was developed to amplify a conserved fragment from within the kinase catalytic domain. Screening an S29 library with this generic SRK probe suggested a minimum copy number of 20 per haploid genome. Six different clones were picked and characterised and their familial relationships established by DNA sequencing. The authentic, S-linked SRK29 gene was isolated by both genomic and cDNA cloning. Its expression and that of the other kinase genes characterised was assayed by reverse transcriptase-

PCR. The SRK appears to be truly stigma-specific. *K3* transcripts were detected in stigma, leaf and root and *K2* in stigma and leaf. The RT-PCR revealed the expression of two new members, *K7* and *K8*, which appear to be leaf- and root-specific. Sequencing of various RT-PCR products revealed loss of a cryptic intron from some SRK transcripts and the persistence of partially processed intermediates of several other kinase mRNAs. RFLP analysis confirmed S-linkage for SRK29 and, possibly, for the *K2* and *K8* genes.

In Durham the complexity of the S-receptor kinase (SRK) class of genes has been further extended with the discovery of a unique variant in the SRK clones selected from the S29 library. The same gene sequence has also been isolated by the Lyon group as a cDNA clone. The gene comprises an atypical S-receptor domain and the last three exons and two introns of the kinase domain (cf. SRK6 gene). Interestingly the coding sequence is truncated with a stop codon but a complete mRNA appears to be transcribed. Both sequences show floral tissue specific expression but are unlikely to represent either the actual SRK sequence and do not appear to be linked to the S-locus. Durham and Lyon are collaborating to elucidate the status of this gene.

Using a kinase domain-specific probe and an SLG5 probe against the same S-locus genomic blots the Durham group has observed that the number of SRK sequences in the S-locus of the S5 line is at least two, i.e. there is more than one SRK5 gene in the S-locus unless both SLG5 and SRK5 are much closer than the distance predicted, or there is a very rare coincidence in the restriction pattern of these two genes.

Pollen-recessive SLG genes, and SLG genes in self-compatible lines

The Lyon group has isolated and sequenced the SLG gene for *Sc* and *Si* lines. The SLG-*Sc* amino-acid sequence deduced from cDNA sequence reveals this glycoprotein to be highly homologous to the pollen recessive S2-allele glycoprotein, whereas the SLG-*Si* predicted protein is homologous to the sequence of SLGs belonging to the S- allele dominant series. We have also determined the sequence of the SRK-*Si* genomic DNA and are now in the process of cloning the SRK-*Sc* gene. A comparison between the SRK-*Si* and -*Sc* gene products should tell us whether the self-compatibility phenotype observed for the *Sc* line is associated with an alteration of the receptor protein kinase activity. At the protein level, an immunochemical analysis has permitted the identification of the SRK proteins of our two *Sc* and *Si* lines. A search for homology between pollen recessive S-alleles (i.e. S2, S5 and S15) and the *Sc*-allele has been carried out by using the molecular probes established from the SLG-*Sc* products. This analysis clearly indicates the existence of homology between the *Sc* and S15 allele. Interestingly, it has also been demonstrated that, in contrast with other findings, the level of expression of the SLG gene is not associated with the strength of self-incompatibility response. This original result strongly questions the importance and role of the SLG gene in the control of self-incompatibility. This work has been undertaken with the collaboration of Dr D. Ockendon (HRI, Wellesbourne).

Nature and operation of S-locus promoters

At Birmingham the expression of a member of the *Brassica oleracea* S-gene family has been investigated in an attempt to determine how S-genes are regulated. The genes involved in the SI response are expressed in a tissue specific and temporally regulated manner, the basis of which must be understood if these genes are to be genetically manipulated to the advantage of the crop breeder. In addition to the commercial application of the work, the S-genes provide a model system for the study of gene regulation in general.

The approach used has been to carry out a deletion analysis of the 5' control region of the gene SLR1 (S-Locus Related gene 1) in order to identify important regulatory elements. This is being followed by the characterisation of the DNA-binding proteins that interact with these elements.

Identification of Regulatory Elements

The SLR1 gene is expressed in the stigma tissue of *Brassica oleracea* buds from approximately one day prior to anthesis; at Birmingham a nested set of fourteen deletions was made in the 5' region of the S63 SLR1 gene and the deleted promoter fragments were linked to the β-glucuronidase (GUS) reporter gene. The SLR1:GUS constructs were transformed into tobacco tissue. The pattern of GUS gene expression was then analysed using both histochemical and fluorometric detection methods.

Results from the deletion analysis revealed that the elements required to direct gene expression in the pistil tissue of transformants are located between 270 and 430 base pairs upstream of the SLR1 gene translation start site. The sequences responsible for the temporal control of SLR1 expression are outside of this region, between the -435 and -1500 position.

Identification and Characterisation of DNA-binding proteins

Work in Birmingham on identifying and characterising DNA-binding proteins that interact with the SLR1 promoter, has been aided by collaboration with the Cologne group of Dr R. Thompson (See 'Major cooperative links'). This was achieved through a short-term EMBO fellowship awarded to Rachel Hackett.

Gel retardation analysis was carried out using tobacco pistil cell extracts. The extracts were found to contain a DNA-binding activity which bound to a motif within an 84 base pair fragment of the SLR1 promoter, covering the region from 279bp to 363bp 5' to the start point of translation.

Work is currently focused on obtaining a cDNA sequence encoding the DNA-binding molecule. A cDNA library has been constructed from tobacco pistil mRNA in the vector λZAPII, which allows expression of cloned inserts. The library is currently being screened using the 84 base pair fragment as a probe.

Identification of S-alleles

An S-allele reference collection of *Brassica oleracea* has been maintained at HRI, and material made available to other members of the programme on request. A molecular method has been developed which enables all 46 S-alleles of this reference collection to be distinguished. The method is based on the amplification of S-sequences using the polymerase chain reaction (PCR) followed by digestion of the amplified products with 6 selected restriction enzymes (PCR-RFLP). The primers for the PCR were designed from a multiple sequence alignment of S-sequence data from published sources and also supplied by M.Trick (Norwich) and C.Scutt (Durham). The restriction enzymes were selected so that S-locus related sequences (SLRs) could be differentiated from putative S-locus glycoprotein (SLG) and S-receptor kinase (SRK) sequences. The technique was tested on various S-homozygotes in a number of crop types and was shown to be consistent.

Analysis of two F_2 populations segregating for different S-alleles showed that there could be problems in using the standard PCR-RFLP technique to identify both the S-alleles present in the heterozygotes. This was shown to be due to preferential amplification of one of the S-alleles. Further analysis of a number of heterozygotes showed that it was not always possible to distinguish the two alleles since restric-

tion patterns are not unique to a particular combination of S-alleles. However, by modifying the PCR-RFLP technique it was possible for the identification of the S-alleles also to be made in particular heterozygotes.

The S-locus in sugar beet

A lambda Zap II cDNA library from sugar beet styles has been constructed at Danisco Biotechnology (Maribo Seeds). This library has been screened with a sugar beet PCR product (SBkin600) which shares identity with the S-genes (SLG and SRK) in *Brassica*. Three clones/plaques hybridized, designated PL1, PL2 and PL3. From the purified positive plaques, plasmids were generated by *in vivo* excision. No hybridization was detected when SBkin600 was hybridized to a cDNA library derived from sugar beet leaves.

The sizes of PL1, PL2 and PL3 were 1.6 kb, 2.9 kb and 2.4 kb, respectively. The entire nucleotide sequence of the three clones has been determined. From the cDNA sequences it can be concluded that, for the first time, it has been proven that an SI-related gene is expressed in sugar beet. From the sequence of PL3 an ORF, more than 634 amino acids in size, can be deduced. This ORF contains the 11 conserved cysteine residues which have been found in all known SLG/SRK amino acid sequences. In addition, the sequence of PL2 shares identity with the kinase domain of SRK. Part of the nucleotide sequence of PL3 has been cloned in a transcription vector, for carrying out *in situ* hybridization on sections of sugar beet styles. Preliminary results indicate that PL3 is expressed in the transmission tract of the style of the sugar beet flower. Collaboration with Dr. R. Thompson (MPI, Cologne) has played a key part in the success of this research.

Transfer of S-alleles to new lines

Transformation of *Brassica oleracea* proved a considerable problem throughout the programme. Despite advice from industry, and a number of 'mini-meetings' this species remained very difficult to transform. Further, the discovery of SRK by the Cornell group at the beginning of the programme strongly suggested that for effective transfer of SI to new lines, recipient plants would have to be transformed with several hundred kb of sequence. For these reasons the Norwich group chose to target and ablate SLR1 function through transformation with an SLR1 antisense construct, using both compatible and self-incompatible recipients. DNA sequence information on the endogenous SLR1 genes of each of these hosts showed high levels of homology, promising a 'universal' antisense effect. Three *B. napus* cv Westar and four *B. oleracea alboglabra* S5 primary transformants were generated. Southern and genetic analysis indicated that each transformant carried a single T-DNA insertion site but it is possible that the *oleracea* plants arose from the same event. Northern blots of stigma RNAs showed both the specificity and effectiveness of the antisense construct. In six of the plants no SLR1-specific message was detectable and one Westar plant showed a greatly reduced level; in each case the endogenous SLG gene appeared to be unaffected. SLR1 ablation was confirmed by Western blotting of IEF gels using SLR1-specific antibody from Lyon. Nevertheless, all of the plants appeared to retain their original phenotypes; the Westar transformants were SC and the S5 plants remained self-incompatible. The pollination behaviour of homozygous T_2 plants selected from each family are now being analysed and transgenic material is also being given to Oxford in order to study possible effects on pollen-stigma protein interactions.

The problems with *Brassica oleracea* transformation led the Birmingham group to perform deletion analysis of promoters in transgenic tobacco. An attempt was made to transform *Brassica oleracea* using different vectors and *Agrobacterium*

strains to try and develop an efficient *Agrobacterium*-based transformation system for *B. oleracea*. Unfortunately, although transformation of tissue from cotyledon petioles was demonstrated, it proved extremely difficult to recover entire transgenic plants from these cells.

The Industrial Perspective

While our understanding of SI in *Brassica* has been considerably advanced, the overall aims of producing a completely reliable SI system, and the means to transfer it to new lines have yet to be achieved. The standards for purity of commercial F1 hybrid *Brassica oleracea* are continuously increasing and, at this moment, the criteria for producing commercially acceptable seedlots are being met. However, since no immediate improvements on the current SI system are on the horizon, at least one company (S&G Seeds) is testing the possibility of producing F1 hybrid vegetable *Brassica*s employing the Ogura cytoplasmic male sterility system. Commercial introduction of the first hybrids based on this production system is scheduled for later this year.

Nevertheless there are a number of disadvantages inherent in male sterility systems and considerable commercial interest remains in developing SI systems of increased reliability.

MAJOR SCIENTIFIC BREAKTHROUGHS (listed under target themes)

Identification and characterisation of the male component of the S-locus

Development of assay for putative male SI determinants (technology).
Identification of PCP7 — a putative male ligand for stigmatically expressed SRK.
Isolation and characterisation of genes encoding the PCP7.
Demonstration of close physical linkage between a PCP gene and an SRK-like gene.
Demonstration of SLG and SRK expression in anthers. Stigma and anther SLG mRNAs are identical.
Identification in anthers of a new S-gene family member unlinked to the S-locus — the SLR3 gene.
Identification and cloning of other functional pollen-coating polypeptides.

Biochemistry and physiology of the SI response

Demonstration that:

(i) level of expression of the SLG gene is not associated with the strength of self-incompatibility response.
(ii) LG-Sc glycoproteins are expressed along the path of pollen tubes, i.e. in the transmitting tissue and ovary.
(iii) SI involves a dephosphorylation step that is inhibited by okadaic acid.
(iv) Strong physiological similarities exist between rejection of intraspecific and interspecific incompatible pollinations.
(v) Strong cell biological and biochemical homologies exist between the SI rejection response and fungal pathogen-host interactions.

Organisation and expression of the S-locus; relationships between S-alleles

Definition of SRK gene family, cloning of SLR1 stigma-specific promoter.
Identification of a further group of undescribed SRK-like genes.
Identification of the SRK polypeptide in P57*Sc* and *Si* stigmas.
Deletion analysis of S-locus gene promoters.

Transfer of S-alleles to new lines
More efficient methods for transforming *Brassica* crops (technology).
Generation of SLR1-null transgenics.

Identification of S-alleles
Development of a PCR-RFLP method for identifying S-alleles (technology).

MAJOR COOPERATIVE LINKS

Collaboration was a very strong feature of the programme both between participants and with the other SI programme organised by Dr. R. Thompson (MPI, Cologne). Collaboration was achieved by means of visits between laboratories, the provision of materials, skills and services, and meetings. There is insufficient space to list the very large number of visits between laboratories. Examples (not all, for reasons of space) of exchanges of materials and expertise and meetings held are set out below.

The programme newsletter SINEWS (see below) played a key role in collaboration both between the two SI programmes and between participants in the *Brassica* project.

The incompatibility 'network' set up under the BRIDGE programme proved very helpful in the organisation of an EC-sponsored course on MOLECULAR AND CELLULAR ASPECTS OF PLANT REPRODUCTION at the University in Nijmegen from 18-22 October 1993. Scientists from different laboratories gave lectures and practical instruction to PhD students and industrial researchers from several Member States.

Exchange of materials and expertise

N.B. Examples only given for ***Nijmegen, Danisco, Durham, Lyon, HRI , Norwich and Nickerson*** because of space constraints; ***Oxford, Birmingham and Zaadunie (S&G Seeds)***, although not mentioned below, were involved in equivalent levels of collaboration.

Nijmegen:
(i) Dr M. Trick, Norwich, provided two cDNA clones and a genomic library.
(ii) Dr T. Gaude, Lyon sent 5 cDNAs complementary to male SLG-like transcripts.
(iii) Dr R.Scott (Leicester) provided the S5 SLG and SLR1 cDNA clones.
(iv) Dr Ockendon provided a method to distinguish between S5 and S29 by means of PCR.
(v) Information on the extraction procedure of the pollen coat proteins provided by Prof H. Dickinson's laboratory (Oxford) was essential for cloning coat polypeptides.

Danisco Biotechnology:
(i) A copy of the cDNA clones PL1, PL2 and PL3 has been transferred to Dr R.Thompson at the MPI (Cologne).

Durham:
(i) Norwich has supplied Durham with an excellent S29 lambda genomic library as well as a number of characterised cDNA probes.
(ii) Durham is actively collaborating with the Oxford group on the characterisation of the PCP genes, the elucidation of their role in SI and their expression.
(iii) Together with the group at HRI, Wellesbourne, Durham is investigating the possible S-linkage of the PCP genes through RFLP analyses.

(iv) Durham and the Lyon group are jointly investigating the role and expression of an unusual SRK-like gene which both laboratories have isolated.

HRI (Wellesbourne):
(i) Seeds have been supplied to various members of the Bridge group.
(ii) Material from Dr T. Gaude, Lyon has been tested for identity of S-alleles.
(iii) DNA from the S-allele reference collection has been sent to Dr R. Croy (Durham).
(iv) Cooperative links have been established with Dr A.M. Chevre and V. Ruffio of INRA, Rennes, France.
(v) Dr M. Beschorner from the Free University, Berlin, Germany has worked at Wellesbourne for 4 weeks and applied the PCR-RFLP method to the study of *B. napus* and *B. campestris.*

Norwich:
(i) Genomic library supplied to University of Durham, SLG cDNA probes to other participants.
(ii) Clones provided to University of Birmingham to enable SLR1 promoter dissection studies.
(iii) Transgenic material transferred to University of Oxford for pollen coat interaction studies.

Lyon:
(i) cDNA sequences or cDNA clones of the *P57Sc* and *P57Si*-SLG have been provided to different laboratories of the BRIDGE programme (Durham, HRI Wellesbourne, Danisco/University of Copenhagen).
(ii) Antibodies (anti-SLG and anti-SLR1, polyclonal or monoclonal) are available for use in our lab.
(iii) Protein blots of sugar beet (from Maribo Seeds) and of rye (Hannover) were sent for immunostaining.
(iv) Other collaborative work with Durham and HRI, Wellesbourne.

Nickerson:
(i) Provided rapid cycling *Brassica* material to Nijmegen.

Meetings and communications

Workshops

There have been three BRIDGE workshops during this period. All of these have been held in conjunction with a second BRIDGE programme focusing on SI in the Solanaceae, coordinated by Dr R.Thompson. The first such meeting was held in April 1991, hosted in Enkhuizen by Zaadunie BV (now S&G Seeds). At this initial meeting all participants presented their proposed research strategies and decisions were made as to the shape and future coordination of the two programmes. Two further collaborative meetings have since taken place. In April 1992 both groups met in Cologne at a workshop hosted by the MPI and organised by Dr R. Thompson. Dr T-H. Kao (Pennsylvania State University) gave the plenary lecture thus placing the EC-supported work in a fully international perspective. The last workshop was hosted by the ENS in Lyon and organised by Dr T. Gaude. The plenary lecturer on this occasion was Dr J. Walker (University of Missouri) who spoke on transmembrane kinases — a subject of current interest to participants from both programmes. Members of the Plant Industrial Platform were invited to attend the second day of this meeting and heard the plenary address, a summing up of the two programmes (presented by Dr R. Thompson and Prof. H. Dickinson) and a presentation on behalf of the Commission by Dr E. Magnien.

Throughout the duration of the programme, small ad hoc meetings have been organised for specific purposes (e.g. to coordinate transformation strategies).

Programme Newsletter — SINEWS

An important outcome of the first workshop in Enkhuizen was the decision to produce SINEWS — a newsletter to which participants from both programmes contributed. SINEWS was published every 6 months and contained research reports from all participants, registers of expertise and materials available for exchange (see above), and lists of training opportunities available within each programme. SINEWS also carried accounts of scientific meetings of general interest and contained the reports made to the Commission at the end of each financial period. So far there have been six issues of SINEWS, one further edition will be produced later in 1994.

PUBLICATIONS

Joint publications

Dickinson, H.G., Crabbe, M.J.C. and Gaude, T. (1992) Sporophytic self-incompatibility systems: S gene products. Internat. Rev. Cytol. **140**, 525-561.

Doughty, J., Scutt, C., Stanchev, B., Croy, R.D. and Dickinson H.G. (1993) Genes encoding *Brassica* pollen coat peptides. International Genetics Congress (Birmingham, U.K.) Poster.

Elleman, C.J., Franklin-Tong, V. and Dickinson, H.G. (1992) Pollination in species with dry stigmas: the nature of the early stigmatic response and the pathway taken by pollen tubes. New Phytol. **121**, 413-424.

Gaude, T., Rougier, M., Heizmann, P.,Ockendon, D.J. and Dumas, C. Expression of the SLG gene is not associated with the self-incompatibility phenotype determinism in the class II S haplotypes of *Brassica oleracea* (submitted).

Guilluy, C-M., Trick, M., Heizmann, P. and Dumas, C. (1991) PCR detection of transcripts homologous to the self-incompatibility gene in anthers of *Brassica*. Theor. Appl. Genet. **82**, 466-472.

R.K.Ruiter, T. Mettenmeyer, H. Meijer, D. van Laarhoven, G.J. van Eldik, J. Doughty, M.M.A. van Herpen, J.A.M. Schrauwen, H.G. Dickinson and G.J. Wullems (1994) Molecular analysis of *Brassica oleracea* pollen coat proteins in relation to anther development (submitted).

Trick, M. and Heizmann, P. (1992) Sporophytic self-incompatibility systems: *Brassica* S gene family. Internat. Rev. Cytol. **140**, 485-524.

Individual publications

Brace, J., Ockendon, D.J. and King, G.J. (1993) Development of a method for the identification of S-alleles in *Brassica oleracea* based on digestion of PCR-amplified DNA with restriction endonucleases. Sex. Plant Reprod. **6**, 133-138

Brace, J., King, G.J. and Ockendon, D.J. (1994) A molecular approach to the identification of S-alleles in *Brassica oleracea*. Sex. Plant Reprod. **7**, 203-209.

Dickinson, H.G. (1993) Pollen dressed for success. Nature **364**, 573-574.

Dickinson, H.G. (1994) Self pollination: simply a social disease? Nature **367**, 517-518.

Dickinson, H.G. (1994a) Dry stigmas, water and self-incompatibility in *Brassica*. Sex. Plant Reprod. (in press).

Dickinson, H.G. and Elleman, C.J. (1994) Pollen hydrodynamics and self-incompatibility in *Brassica oleracea*. In: Pollen-Pistil Interactions and Pollen Tube Growth. Current Topics in Plant Physiology, Am. Soc. Pl. Physiol. (in press).

Doughty, J., Hedderson, F., McCubbin, A. and Dickinson, H.G. (1992) Interactions between a coating-borne peptide of the *Brassica* pollen grain and S(incompatibility)-locus-linked stigmatic glycoproteins. Proc. Natl. Acad. Sci. USA **90**, 467-471.

Doughty, J., McCubbin, A., Hedderson, F., Elleman, C.J. and Dickinson, H.G. (1992) The role of the pollen grain coating in pollination and self incompatibility in *Brassica oleracea*. In: Angiosperm Pollen and Ovules. Eds. E. Ottoviano, D. Mulcahy, M. Sari Gorla and G. Bergamini Mulcahy. Springer Verlag, pp. 55-65.

Elleman, C.J. and Dickinson, H.G. (1992) Pollen-stigma interaction during self-incompatibility in *Brassica oleracea*. In: Genetic control of self-incompatibility and reproductive development in flowering plants. Eds. E.G. Williams, R.B. Knox and A.E. Clarke. Kluwer.
Elleman, C.J. and Dickinson, H.G. (1993) Papillar wall responses during *Brassica oleracea* pollination. International Genetics Congress (Birmingham, U.K.) Poster
Elleman, C.J. and Dickinson, H.G. (1994) The pollen grain coating stimulates key pollination specific responses in the stigmatic papillae of *Brassica oleracea*. (submitted).
Friry, A., Guilluy, C-M., Heizmann, P., Gaude, T. and Dumas, C. (1991) Self-incompatibility products of the male partner in *Brassica oleracea*. In: 'Angiosperm Pollen and Ovules: Basic and Applied Aspects', eds. E. Ottaviano and C.L. Mulcahy, Springer-Verlag, pp. 94-100.
Gaude, T., Friry, A., Heizmann, P., Mariac, C., Rougier, M., Fobis, I. and Dumas, C. (1993) Expression of a self-incompatibility gene in a self-compatible line of *Brassica oleracea*. Plant Cell **5**, 75-86.
Hackett, R.M., Lawrence, M.J. and Franklin, F.C.H. (1992) A *Brassica* S-Locus related gene promoter directs expression in both pollen and pistil of tobacco. Plant J. **2**, 613-617.
Hiscock, S.J. and Dickinson, H.G. (1993) Unilateral incompatibility within the *Brassica oleraceae*; further evidence for the involvement of the self-incompatibility (S)-locus. Theor. Appl. Genet. **86**, 744-753.
Hiscock, S.J., Dewey, F.M., Coleman, J.O.D. and Dickinson, H.G. (1994) An active cutinase from the pollen of *Brassica napus* closely resembles fungal cutinases. Planta **193**, 377-384.
Hiscock, S.J., Doughty, J., Willis, T. and Dickinson, H.G. (1994) Interaction between stigmatic S-glycoproteins and a 7kDa peptide fraction from the pollen of self-compatible *Brassica napus*. Planta (in press).
Kleman-Mariac, C., Rougier, M., Cock, J.M., Gaude, T. and Dumas, C. (1994) SLG proteins are expressed along the path of pollen tubes in *Brassica* pistils. (submitted).
Mariac, C., Rougier, M., Gaude, T. and Dumas, C. (1992) Effects of fixatives on the antigenicity of *Brassica* S-locus specific glycoproteins and rapid immunolocalization by using the tissue print technique. Protoplasma **166**, 223-227.
Ockendon, D.J., Brace, J. and King, G.J. (1993) Identification of S-alleles in *Brassica oleracea*. Seventeenth International Genetics Congress (Birmingham U.K.) (Poster)

Gamete differentiation and fertilization (BIOT CT-900180)

COORDINATOR:
J. van WENT, Wag. Agr. Univ., Wageningen, NL
PARTICIPANTS:
M. CRESTI, Univ. Siena, Siena, I
J. CALLOW, Univ. Birmingham, Birmingham, GB
J. DUNWELL, ICI Seeds, Bracknell, GB
F. HOEKSTRA, Wag. Agr. Univ., Wageningen, NL
M. KREIS, Univ. Paris-sud, Orsay, F
P. OLESEN, Maribu Seed, Copenhagen, DK
M. STEER, Univ. College Dublin, Dublin, IRL

BACKGROUND INFORMATION

Double fertilization in flowering plants, i.e. the fusion of sperm cells with the egg cell and the central cell resp., is cytologically well-documented. The molecular and cell biological principles of gamete differentiation, recognition and fusion, however, are unknown. Techniques are now available for the isolation of male and female gametophytes and gametes of flowering plants, as well as for the *in vitro* fusion of isolated gametes and the *in vitro* culture and regeneration of fusion products. Based on these techniques, work aiming at the identification and study of genes that control major events during gametogenesis, fertilization and early embryogenesis in flowering plants has been initiated.

OBJECTIVES AND PRIMARY APPROACHES

Our long-term goal is to understand the molecular and cellular mechanisms of gamete differentiation, recognition and fusion, with a view to develop methods to manipulate higher plant fertilization for biotechnological purposes. Our short-term objectives are 1. the identification and characterization of female gametophyte and gamete specific gene expression, 2. the identification of specific gamete plasma membrane characteristics which contribute in the process of gamete recognition and fusion.

RESULTS AND DISCUSSION

1. Large numbers of living embryo sacs of *Petunia* hybrida were isolated using a strict enzymatic maceration technique. Total RNA was extracted from the isolated embryo sacs (gametophytic cells) and from ovule derived protoplasts (sporophytic cells). The isolated RNA was used as a template for cDNA synthesis. Prior to cloning in ZAPII, the cDNA was amplified using the Polymerase Chain Reaction. By subtractive hybridization several ovule and embryo sac specific clones could be identified. These clones are in the process of being analyzed. Homology cloning of embryo sac derived cDNA was used to isolate members of potential embryo sac specific protein kinases. A mixture of degenerate oligonucleotides, corresponding to the highly conserved catalytic subdomains VI and VIII of protein kinases sgg/zw3 and MAP/ERK were used as primers for PCR amplification. Seven partial cDNA clones were obtained. Two classes of clones have been characterized and shown to encode peptide sequences similar to the subdomains VI, VII and VIII of the protein kinases *shaggy/zeste-white* 3 and MAP/ERK. *In situ* hybridization experiments are in progress in order to establish the cell type specific expression in the female gametophytes. In various systems MAP and sgg/zw3 kinases

have been associated respectively with mitotic stimulation and establishment of embryo segment polarity. Our results indicate that, in analogy, it might be possible that homologous plant kinases are also involved in developmental processes during embryo sac development and early embryogenesis in flowering plants.

2. In order to develop an *in vitro* gamete fusion system as bio-assay for fertilization events, a method was developed for the isolation of *Petunia* sperm cells. Sperm cell formation in *in vitro* pollen tubes of *Petunia* can be achieved by culturing in medium containing 20% PEG 4000. Using (ultrastructural) immuno-localization it was shown that glycoproteins are present at the surface of *in situ* gametes. *Fucus* gamete-specific monoclonal antibodies did not show any cross reactivity with *Petunia* gametes. Mass isolation of *Petunia* sperm cells is now in development for the production of *Petunia* specific sperm cell surface monoclonal antibodies. A micro-culture method was developed for *in vitro* culturing of individual, isolated *Petunia* embryo sacs and egg cells. Intact embryo sacs and isolated gametophytic cells can be kept alive and metabolically active for at least one week. Cellular analysis of isolated and cultured structures is in progress. The specific distribution of membrane-Calcium and Calmodulin in embryo sacs was determined, using chlortetracycline and fluphenazine resp.

MAJOR SCIENTIFIC BREAKTHROUGHS

Using a RT-PCR clining strategy we were able to identify three genes, which are expressed in the ovules and encoding for putative serine/threonine proteine kinases. With RNA blotting and *in situ* hybridisation it was demonstrated that two of these genes, PSK4 and PSK6 (named PSK for *Petunia* Shaggy-related Kinase) showed clear differential expression in the vegetative and reproductive organs. As expected, both PSK4 and PSK6 are expressed in the ovules. However, while PSK4 is also expressed during vegetative growth, PSK6 transcripts are barely detectable in vegetative tissues. The third gene, PMEK1 (for *Petunia* MAP/ERK-related protein Kinase), appeared to be expressed in both vegetative and female reproductive organs, but not in male organs.

MAJOR COOPERATIVE LINKS

Isolation of *Petunia* embryo sacs, preparation of RNA probes and construction of cDNA was done at Wageningen. Subtractive hybridisation and homology cloning, and characterization of selected clones was done at Paris. Cytology, immunolocalization and *in-situ* hybridization was done at Wageningen. *Fucus* gamete-specific monoclonal antibodies, produced by Birmingham, were tested for cross reactivity on *Petunia* gametes at Wageningen. Methods for the isolation and in vitro fusion of *Petunia* sperm and egg cells are in development at Siena and Wageningen, in collaboration with Hamburg. For this collaborative work and the exchange of methodology considerable exchange of personnel has been organized. Two persons from Paris have been working in Wageningen (8 month in total) for the isolation of embryo sacs and preparation of RNA, three persons from Wageningen have been working in Hamburg (3 month in total) for the isolation and culturing of embryo sacs and gametes, one person from Wageningen has been working in Siena for the development of *Petunia* sperm cell isolation, one person from Dublin has been working in Wageningen on *in vitro* culturing of *Petunia* pollen tubes, one person from Copenhagen has been working in Siena on *in vitro* culturing of pollen tubes.

PUBLICATIONS

Joint publications

Cresti M, Ciampolini F, Van Went J (1991) Strip-shaped projections at the cytoplasmic face of the outer membrane of the generative cell in *Amaryllus belladonna.* Ann Bot **68**:105-107

Decroocq-Ferrant V, Van Went J, Bianchi M, De Vries S, Kreis M (1994) Homologues of *shaggy/zeste-white* 3 protein kinase genes expressed selectively in female and male reproductive organs of *Petunia* hybrida. The Plant Cell (in preparation)

Decroocq-Ferrant V, De Vries S, Van Went J, Kreis M (1994) A member of the MAP/ERK family of protein kinases is preferentially accumulated in female organs of *Petunia* hybrida. The Plant J (in preparation)

Ferrant V, Van Went JL, De Vries S, Kreis M (1992) Construction of PCR cDNA libraries from *Petunia* ovules and isolated embryo sacs. In: Dattee Y (ed) XIIIth Eucarpia Congress on Reproductive Biology and Plant Breeding. Imp Malgogne, Angers, pp. 41-42.

Ferrant V, Van Went JL, Kreis M (1994) Ovule cDNA clones of *Petunia* hybrida L. encoding proteins homologous to MAP and *shaggy/zeste-white* 3 protein kinases. In: Scott, Stead (eds) Molecular and cellular aspects of plant reproduction. Cambridge University Press (in press)

Kreis M, Ferrant V, Van Went J (1993) Embryo sac development: Isolation and characterization of marker sequences and developmental regulatory genes. J Exp Bot **44**:7

Murgia M, Detchepare S, Van Went J, Cresti M (1991) *Brassica napus* pollen development during generative cell and sperm cell formation. Sex Plant Reprod **4**:176-181

Theunis CH, Pierson ES, Cresti M (1991) Isolation of male and female gametes in higher plants. Sex Plant Reprod **4**:145-154

Tirlapur GB, Van Went JL, Cresti M (1993) Visualization of Membrane-calcium and calmodulin in embryo sacs *in situ* and isolated from *Petunia* hybrida and *Nicotiana tabacum.* Ann Bot (in press)

Van Aelst AC, Pierson ES, Van Went JL, Cresti M (1993) Ultrastructural changes of *Arabidopsis thaliana* pollen during final maturation and rehydration. Zygote **1**:173-179

Van Aelst A, Van Went J (1992) Ultrastructural immuno-localization of pectins and glycoproteins in *Arabidopsis thaliana* pollen grains. Protoplasma **168**:14-19

Van der Maas HM, Zaal MACM, De Jong ER, Krens FA, Van Went JL (1993) Isolation of viable egg cells of perennial ryegrass *(Lolium perenne L.).* Protoplasma **173**:86-89

Molecular control of genetic instability in regeneration of crop plants (BIOT CT-900154)

COORDINATOR:
A. KARP, IACR, Long Ashton Research Station, Bristol, GB
PARTICIPANTS:
H. LÖRZ, P.T.H. BROWN, Institut für Allgemeine Botanik, Hamburg, D
M. BUIATTI, Univ. Florence, Firenze, I
F. SALA, Univ. Pavia, Pavia, I
L.P. PIJNACKER, Univ. Groningen, Groningen, NL
Y. HENRY, Univ. Paris-Sud, Orsay, F
P.J. KALTSIKES, Athens Agric. Univ., Athens, GR
G. WERSUHN, Humboldt Univ., Humboldt, D
A. VÁZQUEZ, Univ. Complutense de Madrid, Madrid, ES
M. HUMPHREYS, IGER, WPBS, Aberystwyth, GB)

BACKGROUND INFORMATION

Somaclonal variation has lost much popularity as an area of study. It has been used to produce new varieties but the contribution it has made to plant breeding has been limited. Nevertheless, it is a phenomenon of considerable importance in biotechnology. Tissue culture is an essential component in genetic manipulation. Somaclonal variation remains a largely unavoidable consequence of tissue culture and results in undesired variation in regenerated plants. The production of plants that are not true-to-type reduces the commercial value of tissue culture-derived products and often means that somatic hybrid and transgenic plants cannot be used directly. To improve the efficiency and reliability of plant biotechnological products it is essential that the variation can be controlled.

OBJECTIVES AND PRIMARY APPROACHES

This contract is a concerted action which has enabled the group to keep in communication and most importantly **to focus on how to progress in this area.** Our aims were to study mechanisms of somaclonal variation and to identify where future research should be targeted to determine **why** the variation occurs and how it can be controlled.

RESULTS AND DISCUSSION

In **Bristol**, in a collaborative project with Dr. M.R. Davey at Nottingham University, GB, the cytological instability of wheat in cell suspensions was demonstrated to be related its allohexaploid composition. Greatest stability was found in the diploid wheats and greatest instability in the hexaploid bread wheat. Differences in stability were also observed between species within the ploidy levels. In collaboration with the group of Lörz at Hamburg, a time course study of chromosome variation in 7 cell suspensions of *Triticum aestivum*, derived from anther cultures showed that the cytological behaviour of the lines fit one of two catagories (1) the modal chromosome number remained close to the haploid count of 21, but 1 or 2 chromosomes were lost or gained over time, or (2) the cell underwent doubling followed by aneuploidy. The lines showed variable responses in terms of their ability to regenerate but there appeared to be no clear direct correlation between numerical instability and regeneration capacity. Most

chromosome variation had already occurred by the time the earliest sample were taken, even in those 'elite' lines from which the best regeneration was achieved. C-banding, N-banding and *in situ* hybridisation studies have suggested that chromosome rearrangements and loss are not random with respect to the different wheat chromosomes but, again, there appears to be no clear relationship between such qualitative changes and morphogenic capacity. It is our understanding, now, that to progress in this area emphasis should be placed on investigating (1) cell cycle control *in vitro* compared with *in vivo*, and (2) the organisation of the cytoskeleton in cell cultures differing in morphogenic potential. Preliminary experiments at LARS have indicated that the cortical microtubule organisation of morphogenic cell cultures is quite different from that of non-regenerating cultures and that the nucleus may be de-stabilised through changes in the microtubule organisation. More funds are needed to support further work in this area. In collaboration with Prof. Kaltsikes and Dr. Bebeli, in the Athens Agricultural University, changes in DNA sequences in telomeric heterochromatin were detected in somaclones of rye using *in situ* hybridisation. In collaboration with Jokioinen Plant Breeding Institute, Finland, changes in the distribution of chiasmata were observed in regenerants of 4 inbred and 2 outbred genotypes of rye.

The results obtained in **Pavia** have shown that a foreign gene is rather stably intergrated in the heterologous genome; it remains undisturbed in the genome of tobacco plants even after dedifferentiation of the transformed plant to produce calli and differentiation of these to produce new plants. On the other side, quantitative and qualitative changes have been evidenced in the genome of rice cell suspension cultures and in plants regenerated from them. Molecular evidence has also been given for rearrangement and for extensive amplifications of chloroplast DNA sequences in suspension cultured cells.

In **Paris**, expression of essential genes during wheat somatic embryogenesis has been analysed by inducing disomic, ditelosomic and nulli-tetrasomic cells to undergo embryogenesis from immature embryo culture. The lack of chromosome arm 3 DL suppress somatic embryogenesis and further regeneration capacity from short- and long-term cultures. There is also a strong reducing effect of the lack of 1 AL, 3AL and 3BL. The plants were 'regenerated' through somatic embryo development. Cytological evidence suggests that regeneration through somatic embryogenesis does not ensure normality in chromosome complements. By others in the group, it was also demonstrated that specific rearrangements occur in the mitochondrial genome of wheat depending on the cellular fate of the cultures. Although the mitochondrial genomes of wheat cvs. Chinese Spring and Aquila are similar in organisation, their somatic cultures were reproducible different. A specific region of the mitochondrial genome was missing (or below detectable levels) in cultures of cv Aquila, whilst other regions were specifically amplified in the embryogenic Chinese Spring cultures. The new cytoplasms were maternally inherited and crosses indicated that the mitochondrial reorganisations are under nuclear control.

In **Groningen**, microsporogenesis of a tetraploid potato, *Solanum tuberosum* (+) *S. phureja*, and a hypotetraploid *S. tuberosum* (+) desynaptic mutant has been compared with dihaploid fusion partners. The somatic hybrid had a first meiotic division like tetraploid potato. Abnormal spindle orientation occurred at second meiosis, leading to unreduced gametes at various frequencies. Pollen fertility could not be predicted from the fertility of the fusion partners. Karyokinesis and cytokinesis of leaf protoplasts of *Nicotiana tabacum* (tobacco) and *S. tuberosum* (potato) were compared with those of protoplasts embedded in alginate. Embedd-

ing did not influence the frequency of acytokinesis but the calli became more compact through production of more cells. Acytokinesis decreased through evacuolisation. There exists a close relationship between acytokinesis and cell expansion as a consequence of insufficient wall formation.

Immature embryos of the wheats *Triticum aestivum, T. durum , T. dicoccum* and *T. monococcum were cultured in vitro* on media with various concentrations of 2,4-D and sucrose for 3 days. Media components affected DNA replication from the start of the culture. The frequencies of sister chromatid exchanges (SCEs) were dependent on genotype and were not correlated with the degree of ploidy. They increased after doubling of the concentration of 2,4-D and/or sucrose. The mean numbers were lower than observed in root meristems. Immature embryos of regenerants of *T. aestivum* and *T. durum* demonstrated variable SCE frequencies, which may have been caused by mutations in the parental cell cultures. In the *T. aetivum* embryos the lowest frequencies were found in regenerants obtained from explants with the highest frequencies.

Results obtained in **Florence** include:

a) Genetic variability in tomato somaclones selected for increased or decreased competence for active defense to *Fusarium oxysporum* f.sp. *lycopersici*

RAPD analysis was used to determine the frequency of DNA polymorphism in tomato cell lines from the commercial cv. Red River selected for high (F^+) and low (F^-) recognition of the pathogen or for increased and decreased polysaccharide (FL^+, FL^-) and callose (A^+, A^-) content, all differing in the degree of competence for active defense (Storti et al. 1989 TAG 78: 689-695; Guardiola et al. 1994 TAG in press). Data analysis on somaclonal variability, carried out with a series of casual decamers with a 50% GC content, showed a percentage of polymorphism ranging from 64% for the 'plus' group to 80% for the 'minus' group. Differences between groups were estimated with the use of similarity matrices based on the proportion of shared fragments. The average values of similarity were lower when calculated between genotypes belonging to different groups than those obtained comparing clones belonging to the same group. This was confirmed through ANOVA of similarity coefficients within and between groups. Amplification products correlation analysis, based on the evaluation of the degree of covariation of two specific bands, indicated, moreover, the existence of band(s)specific for clones of different groups. Labelled probes produced from different amplification products from three different primers were used to confirm polymorphism on southern transfers from the corresponding reactions. Preliminary data on Southern hybridizations revealed that three fragments used hybridized to different sequences corresponding, probably, to interspersed multicopy DNA in the tomato genome.

b) Genetic variability in tomato somaclones grown on media supplemented with different auxin/cytokinin ratios and with differences in endogenous hormone levels

Exogenous changes were obtained by growing cell clones all derived from long-term cultures of the same cv. (Red River) on three media with different hormone combinations: 0.4 mg/1 2,4-D and 1mg/1 kinetin (A1); 1 mg/1 2,4-D and 1 mg/1 kinetin (A2); 0.8mg/1 2,4-D and 0.2 mg/1 kinetin (A3). The effect of endogenous variation was analysed through a comparison of habituated and non habituated clones and through the use of cell clones transgenic for *A. tumefaciens* auxin and cytokinin-synthesizing genes in which one of the two pathways was inhibited by gene inactivation.

1. *DNA amplification and methylation.* The results obtained showed two kinds of behaviour, one characteristic of auxotrophic clones, the other of habituated and transgenic cells. Both rDNA multiplicity determined by Southern and dot blot hybridisation and rDNA methylation levels screened through the use of different isochizomers (MspI/HpaII, Sau3A/NdEII) or EcoRII, AvaII, HaeIII methylation-sensitive enzymes, on total DNA of different clones, were shown to be strikingly lower in the first than in the second group of tissues. On the other hand, total methylation levels decreased from the heavily methylated leaf DNA to transgenic, auxotrophic and habituated tissue, the latter displaying the higher digestion with methylation-sensitive enzymes. Changes in exogenous auxin-cytokinin equilibria did not modify this general behaviour although different clones were differently affected. Particularly, differences between culture media were observed on tissues growing on cytokinin rich media (A1), showing higher amplification and methylation levels in some clones. The existence of a cytokinin effect on amplification was moreover supported by experiments carried out on transgenic cells, where higher rDNA multiplicity levels were observed in cells transgenic for the 'ipt' gene from *Agrobacterium* ie. in a cell line where the auxin/cytokinin equilibrium had been artificially changed in favour of cytokinins. Our data thus show both an influence of exogenous auxin/cytokinin treatments and of the endogenous physiological state of the tissues on both phenomena studied.

2. *RAPD analysis.* In the same systems RAPD analysis was carried out with a series of 4 oligonucleotide primers (10-mers) and their combinations. A total of 1068 fragments were visualised across clones grown on A1 medium 1085 on A2 and 1046 on A3, and each primer or combination produced approximately 9 -12 intensely straining fragments from each genotype. The number of polymorphic fragments generated by specific primers varied from 1 to 12. Pairwise comparisons between A1/A2 or A1/A3 or A2/A3 clones with each primer showed a great amount of variation. Moreover, the average values of genetic distances between clones showed a dependance of the extent of polymorphism from the physiological condition of the analysed clones. Particularly, a higher degree of polymorphism was observed on 'A3 clones' the lowest being that shown in 'A2 clones'.
Correlation analysis between polymorphic amplification products measured for each genotype grown on each culture medium, also suggested an influence of growth conditions on the presence of specific bands mutually associated.

3. *RAPD analysis on transgenic Davis tomato cell clones.* RAPD analysis was carried out with 12 oligonucleotide primers (10-mers) on 20 untransformed (control) somaclones, 18 somaclones with higher auxin content (4^-: transgenic for *Agrobacterium* iaaH and iaaM genes) and 18 with higher cytokinin content (1^-: transgenic for *Agrobacterium* iaaH and ipt genes) of the commercial cultivar Davis. The 4^- somaclones showed the presence of polymorphic bands with 3/12 primers used (2 with 70% and 1 with 80% GC content), while the 1- somaclones had an invariant pattern with all the primers, thus confirming that a higher auxin content, both exogenous or endogenous, can affect genome stability. Polymorphism was observed in only one of the control somaclones, namely 5Dbk, with 4/12 primers (2 of which had 70% GC content and 2 50%).

c) *Genetic variability in* Nicotiana tabacum *cell clones subjected or not to abiotic stress (atrazine).* The results obtained in this system show a surprisingly high stability at the molecular level of *Nicotiana tabacum* genotypes. RFLP analysis carried out on a generally variable sequence (the entire rDNA unit) and on an anonymous repeated sequence with three restriction enzymes (EcoRI, HindIII, BamHI) and their combinations did not show any variability. A similar result was

obtained for rDNA through the use of the isoschizomers MspI and HpaII, thus showing the absence of variation in DNA methylation patterns. Only in one RAPD profile out of 12 a single, reproducible polymorphism was observed. Both the repeated sequences analysed seemed to respond to *in vitro* culture conditions and to stress by showing an increase (for the anonymous repeated sequence) or a decrease (for rDNA) in the number of copies in the prescence of the herbicide. Moreover, atrazine-adapted and susceptible clones both showed a high multiplicity variation with respect to leaves, thus confirming that qualitative variation for repeated sequences is a common characteristic of somaclonal variants. ANOVA confirmed a higher variability in the atrazine-selected somaclones than in the susceptible ones for all the probes used in this study, except rDNA. Molecular analysis did not reveal any variation in the atrazine target gene psbA, neither in terms of sequence changes in the active site, nor in the number of copies, thus suggesting that cell tolerance may be due, in our case, to an increase of chloroplasts/chloroplast genomes. This was supported by hybridisation with a probe representing the psbA gene and with an anonymous chloroplast DNA fragment (kindly provided by C. Hartmann, Laboratoire Biologie Moléculaire Végétale, Orsay, France), (Bettini, et al. 1992 Plant Physiol. (life Sci. Adv.) 11: 279-284).

In **Aberystwyth**, plants were regenerated from cell suspensions and suspension derived protoplasts of *Festuca arundinacea* (2n = 6x = 42) and from cell suspensions of a pentaploid hybrid between *L. multiflorum x F. arundinaea* (2n = 5x = 35) in which a set of five homologous chromosomes were marked at the PGI/2 locus by distinct alleles. In both experiments, a direct relationship was found in regenerated plants between time in cell suspension and the number of aberrations at the PGI/2 locus. In these and subsequent published work, chromosomal aberrations were observed which were sometimes, but not always, related to alterations in isozyme activity. In addition over 2000 plantlets were obtained by androgenesis of the same pentaploid hybrid from which we have currently obtained approximately 500 established green plants.

In **Athens**, in triticale, for each line, 15 R_3 families of 20 plants each and a control of 60 plants were tested in the field in a completely randomised design with the individual plant as an experimental unit. The experiment was repeated for two years at the same location. Several agronomic traits were studied. For plant grain yield superior R_3 families, yielding significantly more than the control, were obtained in all lines with frequencies ranging from 0.067 to 0.400 (median 0.067); their frequency was higher in the lines possessing telomeric heterochromatin. The Drira 7RL line with telomeric heterochromatin showed the highest frequency of superior lines. In rye, for each line, varying numbers of R_3 families (5-14) of 20 plants each and a control of 150 plants were planted in the field in a completely randomised design with the individual plant as the experimental unit. The experiment was repeated for two years at the same location. Several agronomic traits were studied, including grain yield. For plant grain yield, superior R_3 families, yielding significantly more than the control, were obtained in all lines except line 5R-- (which contained the least heterochromatin), with frequencies ranging from 0.07 to 0.25; their frequency was higher in the lines possessing telomeric heterochromatin and in particular for the 7R lines.

In **Hamburg**, the results of tissue culture have shown very clearly that many cultures are unstable. Major links between Hamburg and Orsay, as well as Hamburg and LARS showed a range of chromosomal variation in a number of *Triticum aestivum* cultures, derived from anther culture. The response of these cultures varied but there was no indications of any relationship between regeneration

capacity and chromosome instability. Analysis of plants following transformation clearly demonstrated that large differences occurred between transgenic plants in the expression level of the transforming DNA. For example, in cases where there was a large number of transforming DNA molecules, there appeared to be a significant reduction in expression levels. In an attempt to resolve our understanding of culture associated variability, we were, after repeated efforts, able to examine single genes in single protoplasts, and thereby, to examine the stability of such genes as the cultures developed. Finally, in an attempt to investigate the sometimes high level of variability in regenerated *N. tabacum* plants, a species of sibling plants, all derived from 1 single protoplast were examined for alterations in methylation differences. This work is still proceeding.

In **Madrid**, no variation was detected in the case of *S. vavilovi*. Somaclonal variation did not appear to be a frequent event in the case of barley (1%). However, a high frequency of mutant plants was observed in the case of *S. cereale* (50.75%) and this frequency was genotype dependent. In this species we observed a high rate of dominant mutations, the presence of more than one mutation per plant and homozygous mutants as well as a high rate of mutations on particular loci. In some cases transposable elements could be implicated.

MAJOR SCIENTIFIC BREAKTHROUGHS

Hexaploid breadwheat *(Triticum aestivum)* is chromosomally unstable in culture due to its hybrid, polyploid nature. To control its predisposition to instability *in vitro*, work should now be focused on the control of the cell cycle in culture and on aspects of the cytoskeletal organisation.

These have been particularly methodological. In fact, PFGE, PCR and, particularly, the RAPD approach have been shown to be valuable tools to produce qualitative and quantitative evaluations of DNA changes in plant cells. Methods have been worked out for the statistical analysis of changes in the genome of regenerated plants. The fate of these changes in the sexual progeny, as well as in back-crosses planned to restore the original genome can now be conveniently evaluated.

Results obtained in the Plant Molecular Biology lab in Orsay indicate that the accumulation of intermediate transcripts, following transformation (*Agrobacterium* and particle gun) is more intense in calli than in seedlings. These results from wheat calli were corroborated by those obtained with *Oenothera* and maize tissue cultures from other research groups. It has been demonstrated, mostly in wheat and maize, that long term tissue culture induces changes in the mitochondrial genome organisation. Plant tissue culture therefore offers an unique opportunity for creating new cytoplasms.

The data obtained suggest a series of possible conclusions:

a) The extent of genome plasticity is genotype-dependent both at the interspecific (*Nicotiana* and tomato) and at the intraspecific (auxotrophic and autotrophic clones of tomato) levels.
b) The extent of molecular variation also depends on the age of the cultures (see the differences in stability between long-term (Red River) and short-term (Davis) tomato cultures.
c) Genotypes differing for genes affecting the physiological organisation of cells also differ in genetic stability.

d) External conditions affecting physiology and/or causing stress may influence the extent of variation and the fixation of specific levels of amplification and methylation as well as specific RAPD patterns.

Chromosome configurations at meiosis in regenerated pentaploid *Festulolium* hybrids indicated increased interspecific recombination. The possible inclusion of a cell culture phase in a conventional breeding programme to enhance introgression is therefore suggested. More somaclonal variation was observed in polyploid *Festuca* than in diploid *Lolium* species. Plants were regenerated from transformed Festuca protoplasts carrying a selectable gene for hygromycin resistance and a marker gene GUS.

It is concluded that for both the rye and triticale material studied, that *in vitro* culture led quickly to the production of superior lines. Telomeric heterochromatin appeared to affect the release of somaclonal variation, depending on the chromosome arm on which it is borne and the background genotype.

We postulate the possible implication of transposition events in the appearance on variants in *S. cereale* associated with a very high rate of mutation. Consequently, we have a particularly good material with which to study the molecular basis of variations promoted through transposition phenomena.

MAJOR COOPERATIVE LINKS

The coordinator hosted the first meeting at Long Ashton Research Station in June 1991. At this meeting ideas and methodologies were exchanged. Major links are with Hamburg, Paris-Sud and Athens Univ.

F. Sala organised a joint meeting of the participants to this concerted action group at the Isola del Giglio, Italy, in May 1992. On this occasion methodologies were exchanged and results and research programmes were discussed.

A visiting research worker from Spain spent two years at IGER between 1992 — 94 studying somaclonal variation in *Festuca arundinacea*. A PhD student at the University of Liverpool is studying anther culture in the pentaploid hybrid *L. multiflorum x F. arundinacea.*

P.J. Kaltsikes hosted a third meeting in Greece

PUBLICATIONS

Karp, A. (1993). Are you plants normal? Genetic instability in regenerated and transgenic plants. *Agro-Food-Industry Hi-Tech.* **4**, 7-12.

The molecular biology of the cell-to-cell movement of plant viruses in relation to plasmodesmatal function (BIOT CT-900156)

COORDINATOR:
01. R. HULL, John Innes Centre, Norwich, GB
PARTICIPANTS:
02. K. ROBERTS, John Innes Centre, Norwich, GB
03. T. COLBURN, IBMP, Strasbourg, F
04. P. LENNEE, BIOCEM, Aubière, F
05. I. GARCIA-LUQUE, CIB-CSIC, Madrid, E

BACKGROUND INFORMATION

One of the major routes for cell-to-cell communication in plants is via plasmodesmata which are essentially 'dynamic holes' in the cell walls providing symplastic continuity between cells. Since development in plants arises from the differential expression of the plant genome, the movement of macromolecules through plasmodesmata must be controlled. This is achieved by limiting the size of molecules able to pass through them. However, detailed analysis of plasmodesmatal structure, molecular organization and function has been hindered by difficulties in their isolation.

Many plant viruses move from the initially infected cell to other parts of the plant by a combination of short-distance spread through plasmodesmata and long-distance spread via the phloem sieve tubes. The infection unit of a virus is much larger than the normal gating capacity of plasmodesmata and there is increasing evidence that this problem is overcome through the mediation of virus-encoded proteins which facilitate movement from cell to cell. This interaction between virus gene product(s) and plasmodesmata could prove a powerful tool in the understanding of plasmodesmatal structure and function and could be a target for non-conventional protection.

Various lines of evidence indicate that there may be more than one mechanism by which viruses move from cell to cell. In the first, exemplified by *tobacco mosaic virus* (TMV), the movement protein (MP) appears to interact with plasmodesmata only temporally at the infection front allowing movement of the infection unit, which is probably not a complete virus particle, to the next cell; there are no obvious structural changes to the plasmodesmata. In another mechanism, shown by *cowpea mosaic virus* (CPMV), the MP is found associated with highly modified plasmodesmata, which frequently contain virus particles, even at late stages of infection. When the sequences of (putative) MPs are compared they appear to fall into several groups, one based on TMV, one based on *alfalfa mosaic virus* (AMV), one based on CPMV and a loose group of proteins which show little or no relationship in primary amino acid sequence. There is no relationship between the mechanisms described above and the grouping of MPs based on sequence. Similarly, studies on functional complementarity between the movement proteins of several viruses have not identified groupings compatible with those described above.

OBJECTIVES AND PRIMARY APPROACHES

The objectives of this project were to address the following questions:

1. How many apparently different mechanisms are there of cell-to-cell movement of viruses? This question was addressed primarily by participants 01, 03 and 05 who identified the MPs of several viruses using various molecular and cytological approaches. The biochemical properties of the MPs of several viruses expressed in heterologous cell systems was examined. The effects of targetted mutagenesis on the functioning of MPs was studied; this was coupled with the transformation of plants with normal and mutagenized MPs. A system for examining the gating capacity of plasmodesmata by injecting dyes of different molecular radii was developed (participant 02). This was used a) to examine the effects of introducing molecules such as in vitro-produced MPs into cells; b) to examine the effects of mutagenesis of MPs on the gating capacity; and c) to compare the gating capacity in plants supporting replicating virus with that of plants transgenic in the MP. The latter showed if the MP alone can function in fully gating plasmodesmata and also give further information of the differences between transiently and constitutively expressing mechanisms.

2. What are the basic interactions between the virus, MPs and plasmodesmata? Are there receptor proteins and, if so, are they the same for different viruses? Expressed MPs were used in binding studies to isolate possible receptor proteins from hosts (participant 01).

3. What is the physical structure of plasmodesmata and how do receptor proteins fit in? How do plasmodesmata function? Participant 02 used high resolution electron microscopy on thin sections and freeze slam deep-etched replicas coupled with immunogold probing for receptor and/or other host proteins isolated in the binding studies which gave much information on plasmodesmatal structure. Differences in structure induced by the interaction of MPs, coupled with biochemical properties of the host proteins gave leads as to their function.

4. Can the knowledge of interactions between MPs and plasmodesmata be used to design molecules which block viral movement and thereby confer resistance? This was the prime concern of participant 04 who, in collaboration with 05, made plants transgenic in the MP gene of cucumber mosaic virus (CMV).

RESULTS AND DISCUSSION

Major findings on addressing the above questions included:

1. Properties of MPs

a) Identification of the MPs of several viruses and production of antisera.

The infectious genome of beet *necrotic yellow vein virus* (BNYVV) has two RNA species. RNA-1 is infectious to protoplasts but for infection of plants RNA-2 is also required. This indicates that RNA-2 contains the MP gene(s). The sequence of RNA-2 revealed 6 open reading frames (ORFs) and by mutagenesis 3 adjacent ORFs were implicated in cell-to-cell movement. Mutations of the viral coat protein did not interfere with local movement. Thus BNYVV resembles the potex-, carla- and hordei-viruses in having a triple gene block implicated in virus movement.

Grapevine fanleaf virus (GFLV) also has two RNA species and its MP has been identified as a polypeptide product of RNA-2. RNA-2 encodes a 122 kDa polyprotein which is processed into a N-terminal 28 kDa protein of no known function, a 38 kDa protein with amino acid similarity to TMV MP and the C-terminal 56 kDa coat protein; the coat protein is also required for movement. The gene for

the 38 kDa protein has been cloned, expressed as a fusion protein and an antiserum raised against it.

The nucleotide sequence of a Spanish strain of CMV (CMV-24) has been determined. The gene for the putative MP, 3a, was cloned in *E. coli*, expressed and an antiserum raised against it which was used to examine the presence and subcellular localization of the 3a protein in infected and transformed plants. In transgenic and infected plants the 3a protein is detected in the cell wall-enriched fraction as well as in a crude membrane/organelle fraction and a soluble fraction. It was not found in the nucleolar fraction which is in contrast with a previous report, J. Gen. Virol., 1988, **69**:2387) for the location of this protein. The subcellular distribution of CMV 3a protein in transgenic tobacco plants resembled more the situation in transgenic plants for TMV MP than that for AlMV. It remains to be determined whether the differences observed are due to differential modifications of the protein or to different mechanisms of action.

The ability to complement the movement of 3a-defective CMV mutants has been determined using viral transcripts with adjacent 10 amino acid deletions between amino acid positions 124 and 224. The behaviour of these mutants in untransformed and transformed tobacco plants is currently being investigated.

The presence of the CMV 3a protein in the cell wall fraction of transgenic tobacco plants, and its ability to increase the plasmodesmatal size exclusion limit (SEL) (see below), are consistent with the properties of other plant viral MPs, and strongly suggest that the 3a protein is involved in the cell-to-cell movement of CMV.

b) Effect of the MP on the plant cell.

The MP of the GFLV induces tubules in much the same way as do comovirus MPs. The mechanism of movement of nepoviruses is thus like CPMV. The MP of AMV expressed in transgenic tobacco is predominantly found in a cell-wall fraction. It increases the SEL of plasmodesmata from about 800-1000 Da to between 4.4 and 10 kDa without visibly affecting their ultrastructure. Its effect on the SEL is however much weaker than the effect of the TMV MP.

Gene 1 of *cauliflower mosaic virus* (CaMV) which encodes P1 was shown by mutagenesis to encode the MP. Immunofluorescent staining of Chinese cabbage protoplasts infected with CaMV using anti-P1 serum revealed long thread-like structures extending from the protoplast surface. Electron microscopical examination of immunogold-labelled protoplasts showed extracellular tubular structures containing P1.

The presence of the CMV 3a protein in transgenic tobacco plants increases the plasmodesmatal SEL to, in many cases, over 10,000 Da. CMV MP thus causes a functional equivalent modification to those induced by TMV and red clover necrotic mosaic MPs.

In control plants, movement of the largest 10,000 Da probe was not detected. However, in 12% of injections a fluorescent-dextran of 4,400 Da was observed to move into the neighbouring cell. This observation is consistent with previous reports that the SEL under certain physiological conditions is occasionally higher than the commonly reported 800 to 1,000 Da in a variety of tissues.

c) Functional mapping of the AMV MP.

The MP of AMV is the 300-aa non-structural protein (P3) encoded by the 5'-proximal cistron of RNA3. Its gene was introduced into *Nicotiana tabacum* (cv.

Xanthi and Xanthi nc) and *N. benthamiana* behind the 35S promoter. Subcellular fractionation showed P3 to be mainly located in cell walls.

The P3 gene was also expressed in *E. coli* and yeast. The protein was purified and was renatured in the presence of non-ionic detergents. It bound single-stranded RNA or DNA non-specifically in a strong and co-operative manner at low salt but was 80% inhibited by 0.2 M NaCl. The binding was biphasic, with a fast association step and a slow 'stabilization' step; the latter did not occur on ice and might involve a conformational change of the RNA or protein. In contrast, binding of the TMV MP to nucleic acids survives 0.6 M salt and is not temperature-dependent. Altogether, these results show that the properties of P3 and those of the TMV MP are qualitatively similar but quantitatively very different.

To map the different functions of P3, ten in-frame deletion mutants of the gene were engineered and transferred into *N. tabacum* (cv. Xanthi and Xanthi nc). Eight of these, covering the whole coding region, were expressed in *E. coli* and their RNA binding properties tested *in vitro*. In this way an RNA binding domain was identified between amino acids (aa) 36 and 81. In transformed plants, many of the deleted proteins were produced at very low or undetectable levels, being presumably unstable, but those carrying deletions within the N-proximal part of the molecule appeared stable, especially Δ(1-12) (deletion of amino acids 1-12), Δ(21-34) and to a lesser extent Δ(1-77). Of those mutants, only Δ(1-12) fractionated with cell walls at more than 90%. This was confirmed by immuno-electron-microscopy. Δ(1-77) was found predominantly in a cytoplasmic particulate fraction and Δ(21-34) partitioned about equally between cytoplasm and cell walls. This suggested that a domain covered by aa 13 to 77 was involved in cell-wall targeting. The plasmodesmata of plants expressing the three N-proximally deleted proteins had the same SEL as those expressing P3, showing that the N-terminal 'cell-wall targeting domain' was distinct from the 'plasmodesma-modifying domain'. Xanthi nc tobacco expressing Δ(1-34) at high level had a stunted phenotype, abnormally large accumulations of starch in source leaves, frequent necrotic areas, accumulation of acidic PR proteins in the interstitial fluid and other abnormalities suggesting that the transgene was toxic. However, they eventually recovered and were fertile. Xanthi tobacco expressing the same construction at high level looked more 'normal' but produced abortive flowers. The two cultivars differ by the presence (Xanthi nc) or absence (Xanthi) of the N-gene. Whether this is the cause of the phenotypic difference is unknown at the moment.

2. Interaction of MPs with host proteins.

This objective was addressed using TMV P30. A comparison was made of the production and properties of P30 expressed in the baculovirus system, in plants, in *E. coli* and from *in vitro* translation of mRNA transcribed from a cDNA clone. The protein from translation and from plants was produced at low level and that from plants was difficult to purify. *Baculovirus* and *E. coli* expression gave high yields but in both systems the major problem was insolubility of the product. A wide range of treatments was tested in attempts to improve solubility and one of the best methods was to treat the protein first with Sarkosyl and then replace the Sarkosyl with Triton X-100.

Most of the models for MP function suggest that it interacts with one or more host proteins located in plasmodesmata. Since it is difficult to isolate plasmodesmata and characterize them at the molecular level it was considered that the MPs could be used to probe for plasmodesmatal proteins. A wide range of approaches was used in attempts to detect any host proteins to which TMV P30 bound. For instance, extracts of host proteins were electrophoresed into polyacrylamide gels

and blotted onto nitrocellulose membranes. The membranes were then treated with Sarkosyl/Triton solubilized baculovirus-expressed P30, washed and any binding of P30 to proteins transferred from the gel was detected with anti-P30 antiserum. No bands were reliably detected. To enhance the sensitivity of this sandwich hybridization method ^{35}S-labelled P30 was used to probe the nitrocellulose blot. Even this approach did not reliably detect any host protein to which P30 bound. Another approach was based on an electron microscope observation that indicated that plasmodesmata had actin-like fibres associated with them. To see if TMV P30 interacted with actin, various forms of actin, including cytoplasmic 'skeletons' were isolated from plant cells and incubated with labelled P30. No association of P30 with actin could be detected. It thus appears that P30 might not interact directly with any normal plant protein.

3. Structure of plasmodesmata.

The maize root tip was selected a source material for purification of plasmodesmata as it is well characterized, easily obtainable and is rich in plasmodesmata. A method was devised for preparing cell walls which retained embedded plasmodesmata but were virtually free from contaminating cell debris. Briefly, root tips were plasmolysed, snap frozen and ground to a powder. This was then thawed into buffer and passed through a French pressure cell to strip cytoplasm from the wall fragments, which were then washed thoroughly. The resulting material was used for all further procedures.

Antibodies were raised against crude protein extracts from purified walls and monoclonal antibodies were selected on the basis of punctate cell wall staining. While these proved to immunogold label around plasmodesmata, they also recognised other cell components and labeled multiple protein bands in western blots, most prominently at 130 kDa. All of the monoclonals appear to recognise common carbohydrate side-chains.

A sequential biochemical extraction of wall fragments was used to determine the broad chemical composition of plasmodesmatal components. Techniques of negative staining and rotary-shadowed replica production were adapted to allow the effects of the extractions to be assessed with the electron microscope. Light protease treatment left the gross plasmodesmatal form unchanged, indicating that protein is not a major structural component. Heavy proteolysis, on the other hand, occasionally removed the entire plasmodesma, including the surrounding non-proteinaceous collar. It is possible that this treatment disrupts bonds which hold the collar into the wall matrix.

An investigation of the plasmodesmatal collar using specific antibodies and enzymes indicated that it is not a callosic structure as had been previously reported, but is probably composed of another carbohydrate, as yet undetermined. Callose is found in the vicinity of plasmodesmata, but, in the cell wall outside the collar.

Further extractions of wall fragments were conducted using detergents. These confirmed the membranous nature of the desmotubule and the plasma membrane tube, specifically extracting these structures. A small number of proteins was released by these treatments, some of which may originate in the plasmodesmatal core. Interestingly, one of these proteins is the 130 kDa form recognised by our monoclonal antibodies.

These studies have led to a slightly modified model of plasmodesmatal structure in the maize root tip. More importantly, they have introduced novel ways of looking at plasmodesmata in the electron microscope, and reliable methods for extract-

ing specific plasmodesmatal components. Together these provide a strategy for the potential isolation of proteins associated with the membranous plasmodesmatal core.

4. Possible control by interfering with virus movement.

The CMV 3a gene, provided by participant 05, was cloned in the sense and antisense orientations downstream of a duplicated CaMV 35S promoter and a TMV Ω sequence and upstream of the nopaline synthase terminator into a binary vector derived from pTZ18. These chimeric genes were then cloned between the expression cassettes of the neomycin phosphotransferase (NPTII) gene and the β-glucuronidase (GUS) gene into a binary vector derived from PGA 492. Both recombinant binary vectors were introduced into a disarmed *Agrobacterium tumefaciens* strain LBA 4404 and the recombinant strains used to transform tobacco leaf explants. The young regenerated plantlets were selected on kanamycin and screened for GUS activity to check the integrity of the T-DNA insertion into plants. Plants expressing the GUS gene were rooted twice on kanamycin and were transferred to the greenhouse.

Using the serum raised against the CMV 3a protein we were able to detect a 30 kDa protein in the soluble total protein fraction from the transgenic plants transformed with the sense construct. This protein was not detected in control plants expressing NPTII or GUS genes. The 30 kDa protein comigrates with the movement protein found in CMV-infected tobacco tissues, indicating that it was the expected one. Plants expressing the 3a protein were transferred to the glasshouse and seed obtained from those with high, medium and low levels of expression was sent to participant 05 for further characterization.

The expected transcript of about 1300 nucleotides was detected using a specific cDNA probe on total RNA extracted from plants transformed with the 3a gene in the antisense orientation. This RNA transcript was not detected in control plants. Seed from these plants has been sent to partner 05 for challenging the plants with virus particles and viral RNA.

Ten mutated genes (Mx: M1 to M10) from the CMV 3a gene were obtained from participant 05. Each mutant corresponded to a 8 to 12 amino acid deletion between positions 124 to 224. These mutants were cloned into the plant expression vector, transferred to *A. tumefaciens* LBA 4404 and transformed into tobacco leaves as described above. Using the serum against the CMV 3a protein we were able to detect 8 of the mutated proteins (M1 to M6, M8 and M9). These proteins were not detected in control plants. We were not able to detect the M10 protein and are currently checking for the M7 protein. Three of the proteins (M1, M2 and M3) have very different electrophoretic mobilities in comparison with the wild-type MP and protein M2 gave a very low signal which could suggest that it is unstable. These plants are being used to see if the expression of the mutants inhibits the spread of wild-type virus.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

Most of the scientific breakthroughs were in the determination of the properties of the MPs of different viruses and in the refining of the understanding of how the MPs function. The analysis of purified plasmodesmata, although unlikely to throw light on MP function is giving a major advance in the information on plasmodesmatal structure. It seems likely that there could be industrial applicability in the use of defective MPs.

MAJOR COOPERATIVE LINKS

The group held three formal meetings where information was exchanged and cooperation arranged. There were numerous informal meetings at scientific conferences and 10 newletters, View through the Plasmodesma, were issued. There were 6 research exchanges between participating laboratories.

PUBLICATIONS

Joint publications.

Atkins,D., Hull, R., Wells, B., Roberts, K., Moore, P. and Beachy, R.N. (1991). The TMV 30kD movement protein in transgenic tobacco plants is localized to plasmodesmata. *J. Gen. Virol.* **72**: 209-211.

Atkins,D., Roberts, K., Hull, R., Prehoud, C. and Bishop, D.H.L. (1991). Expression of the tobacco mosaic virus movement protein using a baculovirus expression system. *J. Gen. Virol.* **72**: 2831-2835.

Poirson, A., Turner, A.P., Giovane, C., Berna, A., Roberts, K. and Godefroy-Colburn, T. (1993). Effect of the alfalfa mosaic virus movement protein expressed in transgenic plants on the permeability of plasmodesmata. *J. Gen. Virol.* **74**: 2459-2461.

Individual publications.

Erny,C., Schoumacher, F., Jung, C., Gagey, M.-J., Godefroy-Colburn, T., Stussi-Garaud, C., and Berna, A. (1992). An N-proximal sequence of alfalfa mosaic virus movement protein is necessary for association with cell walls in transgenic plants. *J. Gen. Virol.* **73**: 2115-2119.

Gilmer, D., Bouzoubaa, S., Hehn, A., Guilley, H., Richards, K. and Jonard, G. (1992). Efficient cell-to-cell movement of beet necrotic yellow vein virus requires 3' proximal genes located on RNA 2. *Virology:* **189**: 40-47.

Kirchherr, D., Wurch, T., Mesnard, J.M. and Lebeurier, G. (1991). Expression of cauliflower mosaic virus gene 1 in *Saccharomyces cerevisiae. Res. Virol.* **142**: 297-302.

Margis, R., Ritzenthaler, C., Reinbolt, J. and Pinck, M. (1993). Genome organization of grapevine fanleaf nepovirus RNA2 deduced from the 122K polyprotein P2 *in vitro* cleavage products. *J. Gen. Virol.* **74**: 1919-1926.

Maule, A.J., Usmany, M., Wilson, I.G., Boudazin, G., and Vlak, J.M. (1992). Biophysical and biochemical properties of baculovirus-expressed CaMV P1 protein. *Virus Genes* **6**: 5-18.

Ritzenthaler, C., Pinck, M. and Pinck, L. (1994). Grapevine fanleaf nepovirus P38 putative movement protein is a stable final maturation product *in vivo*. (submitted).

Schoumacher, F., Erny, C., Berna, A., Godefroy-Colburn, T. and Stussi-Garaud, C. (1992). Nucleic acid-binding properties of the alfalfa mosaic virus movement protein produced in yeast. *Virology.* **188**: 896-899.

Schoumacher, F., Gagey, M.-J., Maira, M., Stussi-Garaud, C. and Godefroy-Colburn, T. (1992). Binding of RNA by the alfalfa mosaic virus movement protein is bi-phasic. *FEBS Letters* **308**: 231-234.

Schoumacher, F., Giovane, C., Maira, M., Poirson, A., Godefroy-Colburn, T. and Berna, A. (1994). Mapping of the RNA binding site of the alfalfa mosaic virus. (submitted).

Thomas, C.L., Perbal, C. and Maule, A.J. (1993). A mutation of cauliflower mosaic virus gene 1 interferes with virus movement but not virus replication. *Virology* **192**: 415-421.

Viry, M., Serghini, M.A., Hans, F., Ritzenthaler, C., Pinck, M. and Pinck, L. (1993). Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. *J. Gen. Virol.* **74**: 169-174.

Genes required for pathogenicity of bacteria to plants and application of knowledge in biological control of diseases of crops (BIOT CT-900168)

COORDINATOR:
M.J. DANIELS, Sainsbury Lab. Norwich, GB (SL)
PARTICIPANTS:
U. BONAS, Inst. Genbiolog. Forschung, Berlin, D (IGF)
A. VIVIAN, Univ. West of England, Bristol, GB (ISF)
C. BOUCHER, CNRS/INRA, LBMRPM, Toulouse, F (LBMRPM)
M. GUILLON, Calliope, Noguères, F (CALLIOPE)
M. KOKKINIDIS/N. PANOPOULOUS, Inst. Mol. Biol. Biotech., Heraklion, GR (IMBB)

BACKGROUND INFORMATION

Bacterial pathogens cause severe crop losses throughout the world. No satisfactory chemical control methods are available. Disease control can only be achieved by cultural procedures and by breeding disease-resistant plant varieties. Success in both areas will depend on understanding fundamental molecular biology of bacteria-plant interactions.

OBJECTIVES AND PRIMARY APPROACHES

The participants have employed molecular genetic techniques to investigate the following factors required for pathogenicity of bacteria to plants:

i) *rpf* genes of *Xanthomonas* (regulate pathogenicity functions): characterisation of RpfC protein, sequencing of certain *rpf* genes, interaction of RpfN protein with promoters of pathogenicity genes, investigation of a diffusible regulatory molecule, biochemical functions of other *rpf* genes.

ii) *hrp* genes of *Pseudomonas* and *Xanthomonas* (required for pathogenicity and resistance induction): sequencing, characterisation, crystallisation and biochemical function of products, control of expression, role in pathogenicity.

iii) *avr* genes of *Pseudomonas* and *Xanthomonas* (determine inability to colonise certain hosts):
   — sequencing, characterisation of products, role in non-host interactions, structure-function relationships, regulation of expression, evolution of new specificity, role of plasmid transfer.
   — *hrp* mutants of *Pseudomonas solanacearum:* use as biological control agents against bacterial wilt.

RESULTS AND DISCUSSION

A. Regulatory genes

SL: Synthesis of pathogenicity factors (extracellular enzymes and polysaccharide) in *X. campestris* pv. *campestris (Xcc)* is regulated positively by the genes *rpfA-H* and negatively by *rpfN*. Sequencing shows that *rpfC, rpfG* and *rpfH* are members of the two component, sensor-regulator prokaryotic gene family. The genes *rpfF* and *rpfB* have been found to be involved in synthesis of a diffusible substance, which has been purified from bacterial cultures. Enzyme production by *rpfF* mutants is

restored by addition of purified factor. Experiments are in progress to determine the nature of the diffusible factor and its role in regulation. Sequencing of other regulatory genes in the cluster has raised the possibility that pathogenicity is regulated by hitherto undescribed mechanisms.The negative regulatory locus *rpfN* contains a single gene, the product of which is required for binding of a protein to a conserved domain of the promoters of some extracellular enzyme structural genes. It is supposed that this binding is required for down-regulation of enzyme synthesis. Expression of *rpfN* may depend on a σ^{54}-RNA polymerase and may be modulated by other, as yet unknown, genes. Fusions of extracellular enzyme structural genes to reporter genes have been produced to study factors to which regulatory genes respond.

***B. hrp* genes**

LBMRPM: The *hrp* gene cluster of *P. solanacearum (Psol)* has been shown to be organised in a minimum of 6 transcription units, 4 of which are controlled by the positive regulator HrpB, which is encoded within the gene cluster. Expression of these four transcription units is induced in the plant and a synthetic medium allowing their expression *in vitro* has been defined. DNA sequencing of the four transcription units revealed the existence of 20 genes, of which 11 code for proteins homologous to components of the type III secretion pathway which is required for pathogenicity of several animal pathogens of the genera *Yersinia, Shigella* and *Salmonella*. The *hrp* genes in the HrpB-regulated transcription units 1-4 are required for secretion of PopA1, the product of the *popA* gene which is HrpB regulated and located adjacent to the *hrp* cluster. PopA1 has been purified and shown to trigger induction of the hypersensitive response in leaves, but it is not required for pathogenicity on host plants. *popA* may be an avirulence gene since PopA1-responsive *Petunia* genotypes are resistant to infection, contrary to non-responsive genotypes.

IGF/ISV: DNA sequence of the 25kb *hrp* region of *X. campestris* pv. *vesicatoria (Xcv)* revealed 21 *hrp* genes organized in 6 loci. Expression of *hrp* genes is suppressed in complex medium but is induced in the plant and in a defined minimal medium. Subcloning of *Xcv hrp* promoters and primer extension studies revealed a novel sequence motif (PIP box) which might be involved in regulation of transcription. Sequence conservation between most of the predicted Hrp proteins of *Xcv* and *Psol* (LBMRPM) was found, although there is no regulatory gene in the *Xcv hrp* cluster and transcription factor σ^{54} is not required for *hrp* expression in *Xcv*. HrpA1 and HrpB3, 2 of 10 Hrp proteins predicted to be membrane-associated, were visualised in the bacterial membrane with antibodies. In addition to conservation of Hrp proteins between *Xcv*, Psol and other plant pathogens there are striking similarities to *Yersinia, Shigella* and *Salmonella* proteins required for secretion of pathogenicity factors, and it is believed that Hrp proteins of *Xcv* secrete so far unknown virulence factors and elicitors of hypersensitivity.

SL: *Xcc hrp* genes are closely related in structure, function and expression pattern to those of *Xcv* and *Psol*. They are not required for synthesis or secretion of known virulence factors (enzymes and polysaccharide).

IMBB: Attempts have been made to purify (better than 90%) milligram quantities of the HrpR and HrpS proteins of *P. syringae* pv. *phaseolicola* for X-ray structural analysis. When expressed in *E. coli* HrpR and HrpS are insoluble, but a purification protocol was developed and yielded sufficient quantities of very pure protein for crystallographic purposes. The crystallization was not very successful, due to the instability of the proteins after a short period of storage and extensive aggregation.

An attempt was made to co-purify HrpS and HrpR, as experimental data suggest heterodimer formation between HrpS and HrpR occurs which could lead to stabilization. Thioredoxin-Hrp fusions may be more suitable for overexpressed in *E. coli*. The HrpN protein of *Erwinia amylovora* is the prototype of a class of Hrp proteins (harpins) that cause necrosis in plants. A protein encoded by a gene in the *hrp* cluster of *P. syringae* pv. *phaseolicola* may be a 'harpin'. Strong amino acid sequence similarity exists between the published $HrpN_{EA}$ protein and the predicted polypeptide of an open reading frame in a *hrp* locus of *P. syringae* pv. *phaseolicola* that has been genetically characterized. This could give insight into the structure of plant receptors believed to be involved in the hypersensitive response.

C. Avirulence *(avr)* genes

IGF/ISV: Analysis of the *avrBs3-2* gene, a naturally-occurring allele of *avrBs3*, has been completed. *avrBs3-2* induces hypersensitivity in tomato but not in pepper, and is 97% identical to *avrBs3*, representative of an avirulence gene family in *Xanthomonas*. Both genes are plasmid-borne and flanked by 62 bp almost perfect inverted repeats. Similar repeats of unknown function have been found in other members of the *avrBs3* gene family. The specificity of the *Xcv* AvrBs3 protein which governs the interaction with pepper cultivars carrying the *Bs3* resistance gene is determined by the 102 bp direct repeats present in the internal region. Overexpression of AvrBs3 and derivatives with different specifities led to a faster plant response. Specific elicitors of the hypersensitive response have not been identified.

UWE: *Pseudomonas syringae* pv. pisi (Psp) is classified into races according to ability to cause disease on tester cultivars of the host (pea). Two avirulence genes of *Psp* have been sequenced, the race 2 gene *avrPpiA1.R2* and the race 3 gene *avrPpiB1.R3*. Work on plasmid transfer and race specificity has shown that acquisition of unrelated plasmids belonging to the *Pseudomonas* incompatibility group Inc P1 (such as RP4) is associated with changes of race specificity from race 2 and race 6 to a race resembling race 4; in the former case this involves both loss of an avirulence gene A2 (confirmed by molecular analysis) and the gain of a novel avirulence gene, not previously expressed in either race 2 or race 6. Race change is not due to DNA acquisition from the incoming plasmid since strains subsequently cured of the plasmid contain no plasmid sequences. The possibilities are that either the incoming plasmid provides some function that causes the observed race change (perhaps activating a transposable element?) or that the plasmid transfers more readily into the new race, thereby selecting cells which would have to pre-exist as a minor component of the original race. No evidence to support the latter possibility has been obtained. Recent work showed that in addition to substantial deletions of plasmid DNA, the changes also involve (at least in race 2 strains) deletion of a 4-8 kb segment of chromosomal DNA. In collaboration with Wye College and Horticulture Research International (GB) genes of *P. syringae* pv. *phaseolicola* which interact with resistance genes in pea to produce non-host resistance have been studied.

D. Protection of plants with avirulent *hrp* mutants

Calliope: Protective avirulent mutants of *P. solanacearum* have been obtained by insertion into the *hrp* genes of an Omega-KM cassette which is no longer transposable. The presence of these protective mutants within the root system, collar and lower part of the aerial stem protects susceptible tomato from further invasion by a virulent *Psol* strain in the most favourable cases, and limits percentage of wilt-

ing in all cases. Bacteriocins are produced *in planta* by some *Psol* strains but selecting for bacteriocin production does not increase the protective effect of *hrp* mutants. Inoculation procedure is important for ensuring effective protection and the factors that ensure a successful root inoculation by *hrp* mutants have been studied. The French Bio Safety Committee (Commission du Génie Biomoléculaire) authorised small-scale field experiments in naturally infested fields in Guadeloupe.

MAJOR SCIENTIFIC BREAKTHROUGHS

Significant similarities of predicted Hrp proteins of *X.c.* pv. *vesicatoria, P. solanacearum* and *P.s.* pv. *phaseolicola* have been found to *Yersinia, Salmonella* and *Shigella* proteins involved in animal pathogenicity. A *hrp*-dependent protein produced by *P. solanacearum* which elicits plant hypersensitivity has been discovered and purified. A diffusible regulator of enzyme synthesis in *X.c.* pv. *campestris* has been discovered and purified and new regulatory genes have been characterised. Insights have been obtained into both the molecular and 'epidemiological' basis of race change (host specificity) in pathogens. Two-gene avirulence specificity may be involved in non-host resistance in *P.s.* pv. *phaseolicola*. Protection of plants by pretreatment with disabled pathogens has been studied.

MAJOR COOPERATIVE LINKS

The group held regular meetings at which research progress was discussed. Two of the meetings coincided with joint events involving other BRIDGE groups with interests in plant pathogens. Biomaterials and technical information have been exchanged within the group throughout the programme. All the participants have additional research collaborations with other European laboratories outside the BRIDGE programme.

PUBLICATIONS

Joint publications

Arlat, M., Gough, C.L., Boucher C and Daniels, M. (1991). *Xanthomonas campestris* contains a cluster of *hrp* genes related to the larger *hrp* cluster of *Pseudomonas solanacearum*. **Molecular Plant-Microbe Interactions 4**: 593-601.

Individual publications

Daniels, M.J., Barber, C.E., Dow, J.M., Han, B., Liddle, S.A., Newman, M.A., Parker, J.E., Soby, S.D. & Wilson, T.J.G. (1993). Plant and bacterial genes involved in interactions between *Xanthomonas* and crucifers. In: **Advances in Molecular Genetics of Plant-Microbe Interactions 2**: pp. 423-432. (E.W. Nester & D.P.S. Verma, eds.) Kluwer Academic Publications, Dordrecht.

Daniels, M.J., Barber, C.E., Dow, J.M., Han, B., Liddle, S.A., Newman, M.A., Parker, J.E., Soby, S.D. & Wilson, T.J.G. (1993). Bacterial genes requires for pathogenicity: interactions between *Xanthomonas* and crucifers. In: **Mechanisms of Plant Defence Responses**. pp. 53-63 (B. Fritig & M. LeGrand, eds.) Kluwer Academic Publishers, Dordrecht.

Daniels, M.J. (1992). Interaction of *Xanthomonas* with crucifers. **Netherlands Journal of Plant Pathology, 98**: (Suppl. 2) 147-155.

Dow, J.M. & Daniels, M.J. (1994) Pathogenicity determinants and global regulation of pathogenicity of *Xanthomonas campestris* pv. *campestris*. **Current Topics in Microbiology and Immunology** *(J. Dangl, ed.)* Springer, (in press)

Dums, F., Dow, J.M. and Daniels, M.J. (1991). Structural characterization of protein secretion genes of the bacterial phytopathogen *Xanthomonas campestris* pathovar *campestris*: relatedness to secretion systems of other gram-negative bacteria. **Molecular General Genetics 229**: 357-364.

Parker, J.E., Barber, C.E., Fan, M.J. & Daniels, M.J. (1993). Interaction of *Xanthomonas campestris* with *Arabidopsis thaliana*: characterization of a gene from *X.c.* pv. *raphani* that confers avirulence to most *A. thaliana* accessions. **Molecular Plant-Microbe Interactions 6**: 216-224.
Fillingham, A.J., Wood, J., Bevan, J.R., Crute, I.R., Mansfield, J.W., Taylor, J.D. & Vivian, A. (1992). Avirulence genes from *Pseudomonas syringae* pathovars *phaseolicola* and *pisi* conferspecificity towards both host and non-host species. **Physiological and Molecular Plant Pathology 40**: 1-15.
Dangl, J.L., Ritter, C., Gibbon, M.J., Mur, L.A.J., Wood, J.R., Goss, S., Mansfield, J., Taylor, J.D. & Vivian, A. (1992). Functional homologs of the *Arabidopsis RPM1* resistance gene in bean and pea. **The Plant Cell** 4: 1359-1369.
Vivian, A. & Mansfield, J.W. (1993). A proposal for a uniform genetic nomenclature for avirulence genes in phytopathogenic pseudomonads. **Molecular Plant-Microbe Interactions 6**: 9-10.
Bavage, A.D., Vivian, A., Atherton, G.T., Taylor, J.D. & Malik, A.N. (1991) Molecular genetics of *Pseudomonas syringae pv. pisi*: plasmid involvement in cultivar-specific incompatibility. **Journal of General Microbiology 137**: 2231-2239.
Moulton, P.J., Vivian, A., Hunter, P.J. & Taylor, J.D. (1993) Changes in cultivar specificity towards pea can result from the transfer of plasmid RP4 and other incompatibility group P1 replicons to *Pseudomonas syringae* pv *pisi*. **Journal of General Microbiology 139**: 3149-3155.
Wood, J.R., Vivian, A., Jenner, C., Mansfield, J.W. & Taylor, J.D. (1994) Detection of a gene in pea controlling non-host resistance to *Pseudomonas syringae* pv *phaseolicola*. **Molecular Plant-Microbe Interactions** (in press).
Bonas, U. (1994) *hrp* genes of phytopathogenic bacteria. **Current Topics in Microbiology and Immunology** (J. Dangl, ed.) Springer, in press.
Bonas, U., Schulte, R., Fenselau, S., Minsavage, G.V., Staskawicz, B.J. & Stall, R.E. (1991) Isolation of a gene cluster from *Xanthomonas campestris* pv *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. **Molecular Plant-Microbe Interactions 4**: 81-88.
Brown, I., Mansfield, J.W., Irlam, I, Conrads-Strauch, J. & Bonas, U. (1993). Ultrastructure of interactions between *Xanthomonas campestris* pv *vesicatoria* and pepper, including immunocytochemical localisation of extracellular polysaccharides and *AvrBs3* protein. **Molecular Plant-Microbe Interactions 6**: 376-386.
Canteros, B., Minsavage, G., Bonas, U., Pring, D. & Stall, R. (1991) A gene from *Xanthomonas campestris* pv *vesicatoria* that determines avirulence in tomato is related to *avrBs3*. **Molecular Plant-Microbe Interactions 4**: 628-632.
Knoop, V., Staskawicz, B.J. & Bonas, U. (1991) The expression of the avirulence avrBs3 from *Xanthomonas campestris* pv *vesicatoriais* not under the control of *hrp* genes and is independent of plant factors. **Journal of Bacteriology 173**: 7142-7150.
Leite, R.P., Minsavage, G.V., Bonas, U. & Stall, R.E. (1994) Detection and identification of plant pathogenic strains of*Xanthomonas* based on amplification of DNA sequences related to the *hrp* genes of *Xanthomonas campestris* pv. *vesicatoria*. **Applied and Environmental Microbiology 60**: 1068-1077.
Schulte, R. and Bonas, U. (1992). A *Xanthomonas* pathogenicity locus is induced by sucrose and sulfur-containing amino acids. **The Plant Cell 4**: 79-86.
Schulte, R. and Bonas, U. (1992). The expression of the *hrp* gene cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and hypersensitivity on pepper and tomato is plant-inducible. **Journal of Bacteriology 174**: 815-823.
Herbers, K., Conrads-Strauch, J. and Bonas, U. (1992). Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. **Nature 356**: 172-174.
Fenselau, S., Balbo, I. and Bonas, U. (1992). Determinants of pathogenicity in *Xanthomonas campestris* pv. *vesicatoria* are related to proteins involved in secretion in bacterial pathogens of animals. **Molecular Plant-Microbe Interactions 5**: 390-396.
Bonas, U., Conrads-Strauch, J., Fenselau, S., Horns, T., Wengelnik, K. and Schulte, R. (1993). Molecular genetic analysis of *hrp* and avirulence genes of *Xanthomonas campestris* pv. *vesicatoria*. In: **Advances in Molecular Genetics of Plant-Microbe Interactions 2**: pp.275-279. (E. Nester & D.P.S. Verma, eds.) Kluwer Academic Publishers, Dordrecht.

Bonas, U., Conards-Strauch, J. and Balbo, J. (1993). Resistance in tomato to *Xanthomonas campestris* pv. *vesicatoria* is determined by alleles of the pepper-specific avirulence gene avrBs3. **Molecular & General Genetics.** (In press).
Arlat, M., Gough,C.L., Zischek, C., Barbaris, P.A., Trigalet, A. & Boucher, C.A. (1992) Transcriptional organisation and expression of the large *hrp* gene cluster of *Pseudomonas solanacearum.* **Molecular Plant Microbe Interactions 5**: 187-193.
Arlat, M., F. Van Gijsegem, J.C. Pernollet and C.A. Boucher. (1994). PopA1, a protein which induces a hypersensitive-like response on specific petunia genotypes, is secreted via the *hrp* pathway of *Pseudomonas solanacearum.* **EMBO J. 13**: 543-553.
Gough, C.L., Genin, S., Zischek, C. and Boucher, C. (1992). *hrp* genes from *Pseudomonas solanacearum* are homologous to pathogenicity determinants and are conserved among plant pathogenic bacteria. **Molecular Plant-Microbe interactions 5**: 384-389.
Genin, S., Gough, C., Zischek, C. and Boucher, C. (1992). Evidence that the *hrp*B gene encodes a positive regulator of the pathogenicity genes from *Pseudomonas solanacearum.* **Molecular Microbiology 6**: 3065-3076.
Boucher, C.A., Gough, M. and Arlat, M. (1992). Molecular genetics of pathogenicity determinants of *Pseudomonas solanacearum* with special emphasis on *hrp* genes. **Annual Review of Phytopathology 30**: 443-461.
Genin, S., Gough, C.L., Arlat, M., Zischek, C., Van Gijsegem, F., Barberis, P. and Boucher, C. (1992). Involvement of *Pseudmononas solanacearum hrp* genes in the secretin of a bacterial compound which induces a hypersensitive-like response on tobacco. **Advances in Molecular Genetics of Plant-Microbe Interactions 2**: pp. 259-266. (E.W. Nester & D.P. Verma, eds.) Kluwer Academic Publishers, Dordrecht.
Gough, C.L., Genin, S., Lopes, V., Zischek, C. and Boucher, C.A. (1993). Homology between the HrpO protein of *Pseudomonas solanacearum* and bacterial proteins implicated in a signal?? peptide-independent secretion mechanism. **Molecular General Genetics.** (In press).
Rosato, Y.B., Destefano, S.A.L. and Daniels, M.J. (1994). Cloning of a locus involved in pathogenicity and production of extracellular polysaccharide and protease in *Xanthomonas campestris* pv. *campestris*. **FEMS Microbiology Letters 117**: 41-46.
Tang, J.-L., Liu, Y.-N., Barber, C.E., Dow, M.J., Wootton, J.C. and Daniels, M.J. (1991). Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar campestris. **Molecular General Genetics 226**: 409-417.
Van Gijsegem, F., Arlat, M., Genin, S., Gouch, C.L., Zischek, C., Barberis, P. and Boucher, C. (1993). Genes governing the secretion of factors involved in host-bacteria interactions are conserved among animal and plant pathogenic bacteria. In: **Molecular Mechanisms of Bacterial Virulence** (C. Kado, ed.) Kluwer Academic Publishers, Dordrecht. (In press).
Van Gijsegem, F., Genin, S. and Boucher, C.A. (1993). Conservation of secretion pathways for pathogenicity determinants of plant and animal bacteria. **Trends in Microbiology 1**: 175-180.
Van Gijsegem, F., Zischek, F., Gouch, C., Genin, S and Boucher, C.A. (1994). Completion of the sequence of the part of the *P. solanacearum hrp* gene cluster required for protein secretion: additional homologies to type III secretion system of mammalian pathogens and to flagellar genes. (Submitted).
Zischek, C., Gouch, C., Niqueux, E., Arlat, M., Van Gijsegem, F., Barberis, P., German, S. and Boucher, C.A. (1994). The *hrp* gene cluster of *P. solanacearum* codes for exported and membrane associated proteins homologous to proteins controlling secretion of animal pathogenicity determinants or biogenesis of the bacterial flagellum. (submitted).

Molecular basis of signalling in *Rhizobium meliloti-Medicago* interactions and genetic improvement of nodulation ability (BIOT CT-900159)

COORDINATOR:
01: A. KONDOROSI, Inst. des Sciences Végétales CNRS, Gif/Yvette, F
PARTICIPANTS:
02: M. JOHN, Max-Planck-Institut für Züchtungsforschung, Köln, D
03: J. OLIVARES, Estación Experimental del Zaidin CSIC, Granada, E
04: J. HACHE, Bertin and Cie, Plaisir, F

BACKGROUND INFORMATION

Nitrogen is a basic element of organic materials. Though, it is present in unlimited amounts in the atmosphere as N_2, this form cannot be utilised by living organisms except a few microorganisms. The most pronounced way of the biological nitrogen fixation is provided by the *Rhizobium* bacteria which, in symbiosis with their leguminous host plants, induce the formation of a unique plant organ, the root nodule which serves as a niche for the nitrogen-fixing form of the bacteria. Establishment of nitrogen-fixing symbiosis is governed by signal exchanges between the host plant and the bacterium that lead to the activation of genes in the symbiotic partner. Nodule development consists of several stages determined by different sets of genes both in the micro- and macrosymbiont. First, flavonoid signal molecules exuded by the host plant roots induce the expression of the nodulation (*nod, nol*) genes in *Rhizobium* in conjunction with the bacterial activator NodD protein. Then, lipooligosaccharide Nod factors carrying various host-specific decorations are produced by the bacterial Nod proteins. The Nod factors induce a variety of plant reactions, such as root hair deformation, initiation of nodule meristems and induction of early nodulin genes leading to nodule formation. Other bacterial genes provide competitiveness of the bacteria for nodulation or are required for successful infection and for nitrogen fixation. Furthermore, for the successful establishment of a symbiotic interaction the induction of plant defense mechanisms has to be avoided.

The aim of the work carried out by the collaborating groups was to elucidate the early communication steps between the micro- and macrosymbiont, by studying the *Medicago-Rhizobium meliloti* interactions. At the start of the project most of the *nod* genes had been isolated, however, their functions were not known, their regulation was not fully understood, the structure of the first Nod signal was just reported and no plant genes involved in nodule organogenesis were isolated.

OBJECTIVES AND PRIMARY APPROACHES

The goals on the bacterium side were understanding

i) the complexity of *nod* gene regulation,

ii) identify Nod factors from *R. meliloti*,

iii) to determine the role of *nod* genes in Nod signal production, and

iv) to characterize bacterial genes providing higher competitiveness for nodulation ability. On the plant side, as a long term project, the objectives were

v) to identify and characterize genes whose expression is induced at very early stages of symbiosis, especially those that are specifically expressed in response

to purified Nod signals and to understand their role in nodule organogenesis. Another goal was

vi) to study and compare plant responses induced by the symbiotic and pathogenic signal molecules by analysing expression of marker genes associated with plant defense.

RESULTS AND DISCUSSION

i) 01: Involvement of Nod regulatory proteins of *R. meliloti* in *nod* gene activation

In *R. meliloti* a low and finely tuned level of *nod* gene expression is required for optimal nodulation. It is achieved by a complex regulatory circuit which involves several activators, differing in their flavonoid specificity, and the NolR repressor displaying a differential control on the *nod* genes involved in the synthesis of the core Nod factor structure. Positive and negative elements controlling the expression of the *R. meliloti* nodulation genes were identified and characterized in detail. Negative regulation by the repressor protein NolR is thought to play a role in the fine tuning of Nod signal production in order to provide optimal concentrations of signals adapted to specific host plant genotypes. Alternatively, a combination of positive and negative regulation may differentially influence the expression of different subsets of *nod* genes. The *nod* genes are organized into several transcriptional units preceded by conserved *nod* box elements, however only some of the operons are subject to downregulation by the NolR repressor. This opens the possibility to alter the balance of *nod* gene products having different functions in Nod factor biosynthesis and in this way to change the ratio of differently substituted Nod signal molecules. The construction of *R. meliloti* strains containing multiple copies of the positive regulatory genes *nodD3* and *syrM* resulted in a high level of flavonoid independent *nod* gene expression that broke the host specific barrier by extending the nodulation ability to the non-host plant siratro. This might be one way to improve biological nitrogen fixation.

ii) 01: Identification of the *R. meliloti* Nod factor family

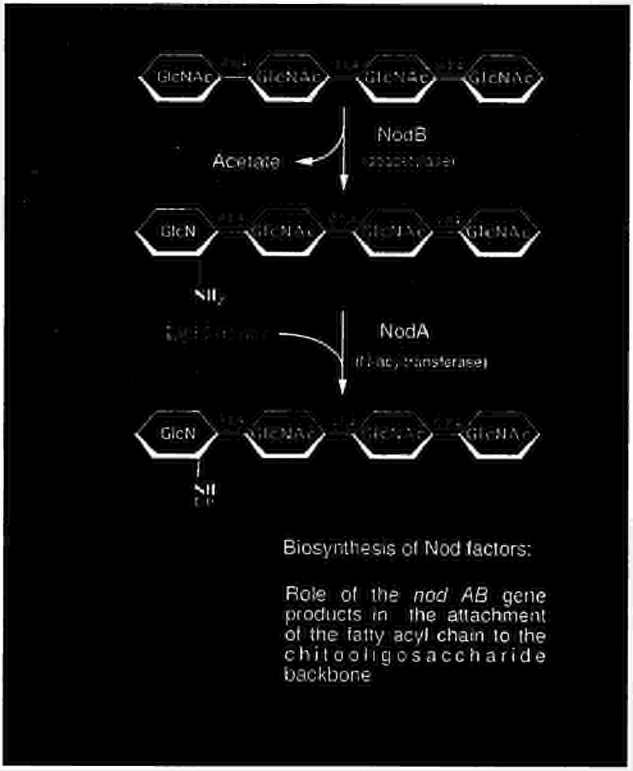
Using our knowledge about the regulatory factors (see i 01) has allowed the manipulation of *R. meliloti* strains lacking the negatively acting factors but containing higher amounts of positive transacting factors. These strains produced more than thousand fold higher quantitites of Nod signals in comparison to the wild type strains. This engineering step was a prerequisite to purify and characterize the Nod metabolites from *R. meliloti*. We demonstrated that a large number of different structurally related Nod metabolites is produced by a single *Rhizobium* species and the major products were identified as tetra- and pentamers of β-1,4-*N*-acetylglucosamine sulfated at the reducing terminus and *N*-acylated with a C16:2 or C16:3 acyl chain at the nonreducing terminus.The ability of the Nod factors to induce morphological and developmental changes on plant roots, like root hair deformation and induction of nodule structures were investigated. The differential activity on host versus non-host plants of related lipo-oligosaccharides allowed to identify some of the structural features (e. g. oligosaccharide chain length) influencing the host plant-specific recognition. The ratio of Nod factors displaying different degrees of activity on a given plant species or genotype is an important determinant of host-specific recognition.

iii) Role of *Rhizobium nod* gene products in Nod factor biosynthesis

***02: NodB is a deacetylase and NodA is an N*-acyltransferase**

The *Rhizobium* nodulation genes *nodABC* are involved in the synthesis of lipooligosaccharide symbiotic signals, which are mono-*N*-acylated

chitooligosaccharides. These bacterial signals elicit nodule organogenesis in roots of legumes. Recently we have shown that the NodA and NodB proteins are sufficient to produce small, heat-stable compounds that stimulate mitosis in various protoplasts derived from legumes and non-legumes. To test whether these gene products could play a role in regulation of plant growth and development, we introduced and expressed the *Rhizobium meliloti nodA* and *nodB* genes singly or in combination under the control of diverse promoters in tobacco. The resulting transgenic plants showed characteristic morphological abnormalities, indicating that substrate molecules are likely to be present in tobacco that allow the *nodA* and *nodB* encoded proteins to synthesize growth controlling factors and also signal transduction pathways to respond to the presence of these regulatory molecules.



To explain the effects caused by these *nod* genes in non-legume transgenic plants, and to elucidate their specific role in the pathway of Nod signal synthesis, we have determined their biochemical function by direct assays. Since the *nodB* transgenic tobacco plants showed the most pronounced phenotype, the function of NodB was investigated first. The NodB protein has been overproduced in *E. coli* and was purified by immunoaffinity chromatography or by renaturation from inclusion bodies. Our data showed that the NodB protein of *R. meliloti* deacetylates the non-reducing *N*-acetylglucosamine residue of chitooligosaccharides. This step provides a free amino group at the non-reducing terminus of the oligosaccharide backbone, which is a necessary prerequisite for fatty *N*-acylation in the pathway of Nod factor synthesis.

The nucleotide sequence of the *Rhizobium meliloti nod* genes revealed that the *nodA* stop codon overlaps the *nodB* start codon suggesting a functional coupling between the two genes. It seemed reasonable to assume that the *N*-acylated glucosamine oligosaccharide signals might be synthesized by the enzymatic acylation of a free amino group released by NodB at the non-reducing terminus of the oligosaccharide backbone. To demonstrate that acyl group transfer is catalyzed by NodA, we synthesized a radiolabelled tetrasaccharide precursor carrying an amino group as a potential attachment site for *N*-acylation at the non-reducing glucosamine residue. An acyl transfer reaction was carried out using this specific substrate as acyl acceptor and rhizobial cell extracts as a source of enzyme and lipid donors. Our data show, that cell extracts containing the NodA protein were able to convert this tetrasaccharide precursor into a butanol-extractable hydrophobic compound that co-migrated on the TLC plate with *R. meliloti* lipooligosaccharide. Moreover, this NodA conversion product migrated on a silica gel plate at a rate comparable to that of a tetrasaccharide bearing a palmitoyl substituent at the terminal glucosamine residue. The hydrophobic NodA product was retained by a C_{18} reversed-phase HPLC column and exhibited the same retention time as non-sulfated *R. meliloti* nodulation factor NodRm-IV (C16:2). To demonstrate that the hydrophobic product has the characteristic β-1,4 linkage of lipooligosaccharides, the HPLC-purified compound was digested with chitinase and yielded the expected degradation pattern. Finally, the HPLC fraction co-eluting with NodRm-IV was shown to elicit deformation of root hairs on vetch, which can be observed only with *N*-acylated chitooligosaccharides carrying a correct structure of the lipid moiety. Thus, various criteria demonstrate that NodA is involved in the incorporation of a fatty acyl chain into the synthetic tetrasaccharide acceptor, to yield biologically active nodulation factor.

01: NodM is a glucosamine synthase providing precursor for Nod factor biosynthesis
We showed that the NodM gene coding for a glucosamine synthase is necessary for a high level of Nod factor production. In addition, this gene contributes to the accuracy of the synthesis of the cognate Nod signals over various analogues. This suggests that NodM, but not its housekeeping analogue, is able to channel precursor molecules to the specific Nod factor biosynthetic machinery.

iv) 03: Characterization of *nfe* genes providing higher competitiveness of rhizobia for nodulation

Legume productivity in agricultural fields may be improved by inoculation with selected highly effective nitrogen-fixing root nodule bacteria. However, field inoculations have often been unsuccessful due to the presence of highly competitive native strains in the soil that compete with the introduced strain in nodule formation on the host plant. Aiming at the characterization of genetic determinants of competitiveness a 20 kb DNA region named nfe providing higher nodule formation efficiency and competitiveness of *R. meliloti* GR4 was located on the megaplasmid pRmeGR4b.

Sequence analysis of the *nfe* region revealed several open reading frames (ORFs). Two of them, *nfeA* and *nfeB*, were preceded by functional *nif* consensus sequences and NifA-binding motifs, though an additional, NifA-independent, transcriptional start site for *nfeA* was also found. A high degree of homology was found between the amino terminal domain of NfeA and NifH. Polyclonal antibodies raised against NfeA revealed that this protein is widespread in rhizobia. 128 base-pairs downstream of *nfeB* a 970 bp long open reading frame was found whose deduced amino acid sequence showed similarity to ornithine cyclodeaminase (OCD) of *Agrobacterium tumefaciens* an unusual enzyme that converts ornithine into proline.

Southern hybridization indicates that strain GR4 carries a single copy of the *ocd*-like gene that was not conserved in other *R. meliloti* strains or *Rhizobium* spp tested. Mutation in the ORF resulted in impaired nodulation efficiency. Therefore this locus was considered as a novel *nfe* gene termed *nfeC*. Downstream of the *nfeC* gene an IS element homologous to *ISRm3* was found. In a 2.6 kb distance from the *nfe* genes, an open reading frame (ORF1) with a coding capacity for a 23-kDa protein was homologous to the putative transposase of *Pseudomonas capacia* insertion sequence IS402. Since homologous sequences have not been described from rhizobia, this IS-like element of *R. meliloti* GR4 was named *ISRm4*. Although this IS element is present as a single copy in strain GR4 novel reiterated *ISRm4* sequences were observed in GR4 derivatives indicating that this IS was effective in transposition. We found that *ISRm3* and *ISRm4* were widespread in the *R. meliloti* strains though the latter was less common and their homologs were found also in other rhizobia.

Divergently oriented from the *nfe* genes two ORFs (ORFA and ORFB) were identified. Though, *nif* consensus sequences were not found, expression of ORFA might be indirectly coupled to the NifA-NtrA regulatory network. The putative ORFB-encoded protein contained a helix-turn-helix motif, that resembles the DNA-binding consensus sequence proposed for many prokaryotic regulatory proteins.

To transfer *nfe* genes into other rhizobia and to integrate them into the chromosome, cloning vectors were constructed using the *oriV* of one of the cryptic plasmids of *R. meliloti* GR4. pJMB45 carried *oriV* on a 4.8 kb *PstI* fragment where six open reading frames were identified. One of them was a homolog of the *repC* found on the Ti and Ri plasmids of *Agrobacterium* strains. The other five ORFs are likely not necessary for replication and stability of the vectors. No conservation of sequences corresponding to other reported replication origins have been found, although the product of one of them, ORF2, showed certain homology to cytoskeletal proteins of the eukaryotic cell.

v) 01 and 04: *Medicago* genes involved in root nodule organogenesis

During nodule development several sets of genes with both specific and general functions are induced e.g.

i) early nodulin genes that are nodule-specific and probably required for nodule structure and function,

ii) genes that have general functions in plant differentiation and organogenesis such as cell cycle genes and

iii) genes whose products interact specifically with the Nod factor and are involved either in its recognition or in its degradation.

We identified the *Enod12A* and *B* genes as well as the *Enod40* genes from *Medicago*. We showed that expression of the *Enod12A* and *B* genes, encoding putative hydroxyproline-rich cell wall proteins, are differentially regulated; *MsEnod12B* is inducible by the Nod factor whereas expression of *MsEnod12A* is detectable only in the differentiated nodules and requires bacterial infection. The finding that mutant plants lacking *Enod12* genes are fully competent for the development of N_2-fixing nodules (Csanádi et al., 1994) may indicate that the gene function is either redundant or is required only under yet unknown conditions. The *Enod40* gene is expressed in all cells of the nodule primordia, then in the differentiating nodule cells, and is uncoupled of bacterial infection. It codes for a non-translated cytoplasmic RNA and appears to be involved in cellular differentiaton. Ectopic expression of the sense and antisense constructs in transgenic *Medicago*

plants affected plant growth and differentiation drastically suggesting similar role of the *Enod40* RNA as the recently discovered mammalian riboregulators. We showed that the active Nod factors but not their inactive derivatives induced differential expression of cell cycle marker genes, stimulated cell cycle progression and and reactivated most likely the G0 differentiated cells for division. All of these genes were mapped in the *Medicago* genome. For this work we could apply all techniques and equipments that were developed by participant 04 during the grant period. This included a new method for DNA extraction and the use the automatic MultiBlotter System (Labimed).

vi) 01: Plant responses induced by pathogenic signals and by the symbiotic Nod factor

In order to get an insight into plant defense responses in *Medicago*, first, leaf reactions were analysed in compatible and incompatible interactions induced by infiltration of leaves with *Xanthomonas campestris* pv. *alfalfae* and *Pseudomonas syringae* pv. *pisi*, respectively. Moreover, they were compared to responses elicited by non-pathogenic bacteria, *R. meliloti* or the non-host *R. leguminosarum* (Rl) or *Escherichia coli* (Ec). None of the non-pathogenic bacteria induced a hypersensitive response (HR) and their growth characteristics inside the leaf were different from that of the pathogenic ones: the number of bacteria remained more or less constant, or slightly decreased within 8 days, but thereafter diminished abruptly and disappeared in 2 days. At the gene expression level, the accumulation of chalcone synthase (CHS) and isoflavone reductase (IFR) transcripts, involved in the production of isoflavonoids and the phytoalexin medicarpin, was analyzed. Maximum accumulation of these transcripts was observed between 45 and 90 min., whereas in a defense reaction induced by the incompatible bacterium *P. syringae* pv. *pisi* the maximum occured 6h postinfection. HPLC analysis of phenylpropanoids suggested that a mechanism for the phenylpropanoid gene induction exists also in non-pathogenic interactions but it might be distinct from that involved in a defense reaction.

The expression of genes involved in isoflavonoid synthesis was also inducible by Nod factors, but only at highly elevated concentrations. The response was much stronger in *Medicago* cell cultures than in roots and correlated with the biological activity of the Nod factors. In contrast, these genes were inducible by low concentrations of unmodified chitooligosaccharides in the cell cultures suggesting that the structural modification brought about by acylation of chitooligosaccharides may render an elicitor of plant defense reactions into a symbiotic signal molecule.

The intriguing similarities between Nod factors and chitin prompted us to investigate whether legume chitinases would hydrolyse Nod factors as well as non-modified chitooligosaccharides. We indeed found that Nod factors are substrates for plant chitinases but were more resistant to hydrolysis by chitinases than the unmodified chitooligosaccharides. Kinetic analysis showed that the structural parameters determining host specificity, the length of the oligosaccharide chain, the acylation at the non-reducing end and the sulfatation at the reducing end of the lipooligosaccharide, influence the stability of the molecule against degradation by chitinases. These results open the possibility that the activity of Nod factors on *Medicago* may partly be determined by the action of chitinases. It is tempting to speculate that in the rhizosphere a higher stability of Nod factors compensates for the lower amounts excreted by bacteria. Moreover, degradation of the Nod factors by plant hydrolytic enzymes may play an autoregulatory role in nodule initiation.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

Better understanding of *nod* gene regulation (e.g. host plant flavonoid-independent *nod* gene expression) resulted in host range extension. Elucidation of molecular events in nodule organogenesis and the function of bacterial and plant gene products increased our knowledge on the molecular and cellular mechanisms of plant differentiation and development. Moreover, construction of transgenic plants expressing either *nod* genes or plant genes involved in nodule organogenesis may be a first step for the creation of a niche for rhizobia to establish symbiosis with non-host, even monocot plants.

Genetic engineering of the *Rhizobium* strains (eg. with genes providing higher competitiveness or extension of host range) as well as that of the host plant might have a significant impact on plant breeding and agricultural yields.

An automatic multisouthern/northern electrophoresis and blotting system has been developed for the establishment of RFLP maps or genetic fingerprints and for gene expression studies.

MAJOR COOPERATIVE LINKS

ISV, MPI and EEZ as well as ISV and BERTIN had regular exchanges of information, plasmids, bacteria and various biological materials. A. Savouré full time and B. Hoffmann for 2 years from BERTIN worked at ISV. With other laboratories participating in the BRIDGE Integrated Action, with Univ. Bielefeld (Prof. A. Pühler): One week stay of two Spanish scientists and a three month stay of a fellow in Bielefeld and two German scientists in Granada.

20-22 Jan. 1991	Meeting of all participants of the BRIDGE project at ISV in Gif.
26-29 Aug. 1991	Visit of Dr. E. Kondorosi (ISV) at MPI for discussion of results, collaborations and exchange of strains.
7 Oct. 1991	Visit of Prof. H. Van Onckelen (Univ. Antwerp, Participant of other BRIDGE project) at MPI for discussion of results, collaboration and collection of samples.
27 Oct. 1991	BRIDGE ELWW meeting on '*Rhizobium*-legume symbiosis' in Capri.
5-8 Nov. 1991	Visit of Dr. M. Schultze (ISV) at MPI for discussions.
24 Feb. 1992	Visit of Prof. H. Van Onckelen at MPI for discussion and collection of samples.
9-10 Mar. 1992	Visit of Drs. H. Röhrig and J. Schmidt (MPI) at ISV for discussion of results, cooperation and exchange of material.
14-17 Mar. 1992	Meeting of all participants of the BRIDGE project at EEZ.
30-31 Mar. 1992	Visit of Drs. M. John and J. Schmidt (MPI) at ISV for discussion of joint projects and exchange of material.
8 Dec. 1992	Award of the Max-Planck Research Prize to Drs. A. Kondorosi (ISV) and J. Schell (MPI) by the Alexander von Humboldt Foundation in Bonn.
17-19 Dec. 1992	Visit of Drs. A. and E. Kondorosi (ISV) at MPI for discussion of results.

26 Mar. 1993	Meeting of all participants of the BRIDGE project at Bertin in Plaisir.
11-12 Apr. 1994	Bridge Final Sectorial meeting 'Plant-Microbe interactions' at Dourdan
20-25 June 1994	Meeting of participants (ISV, MPI) discussing results
29-31 Aug. 1994	Meeting of participants (ISV, MPI, EEZ) discussing results and preparation of manuscripts

PUBLICATIONS

Joint publications

Kondorosi, E., Pierre, M., Cren, M., Haumann, U., Buiré, M., Hoffmann, B., Schell, J. and Kondorosi, A. (1991) Identification of NolR, a negative transacting factor controlling the *nod* regulon in *Rhizobium meliloti. J. Mol. Biol.,* **222**, 885-896.

Prinsen E., Chauvaux N., Schmidt J., John M., Wieneke U., De Greef J., Schell J. and Van Onckelen H. (1991) Stimulation of indole-3-acetic acid production in *Rhizobium* by flavonoids. *FEBS Lett.* **282**, 53-55.

Schmidt, J., John, M., Kondorosi, E., Kondorosi, A. and Schell, J. (1991) Recent progress in the elucidation of the function of *nodAB* and *C* genes of *Rhizobium meliloti. Isr. J. Bot.,* **40**, 165-169.

Kondorosi, A., Kondorosi, E., John, M., Schmidt, J. and Schell, J. (1991) The role of nodulation genes in bacterium-plant communication. In: *Genetic Engineering*, Vol. **13**, (Ed. Setlow J.K.), Plenum Press, New York and London, pp. 115-136.

Grosskopf, E., Cam Ha, D. T., Wingender, R., Röhrig, H., Szecsi, J., Kondorosi, E., Schell, J., Kondorosi, A. (1993) Enhanced levels of chalcone synthase in alfalfa nodules induced by a Fix- mutant of *Rhizobium. Molec. Plant-Microbe Interactions* **6**, 173-181.

Röhrig, H., Schmidt, J., Wieneke, U., Kondorosi, E., Barlier, I., Schell, J. and John, M. (1994) Biosynthesis of lipooligosaccharide nodulation factors: *Rhizobium* NodA protein is involved in *N*-acylation of the chitooligo-saccharide backbone. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3122-3126.

Individual publications/Patents

Györgypal, Z., Kondorosi, E. and Kondorosi, A. (1991) Diverse signal sensitivity of NodD protein homologs from narrow and broad host range rhizobia. *Mol. Plant-Microbe Interact.* **4**, 356-364.

Kondorosi, E., Buiré, M., Cren, M., Iyer, N., Hoffmann, B. and Kondorosi, A. (1991) Involvement of the *syrM* and *nodD3* genes of *Rhizobium meliloti* in *nod* gene activation and in optimal nodulation of the plant host. *Mol. Microbiol.* **5**, 3035-3048.

Baev, N. and Kondorosi, A. (1992) Nucleotide sequence of the *Rhizobium meliloti nodL* gene located in locus n5 of the *nod* regulon. *Plant Mol. Biol.* **18**, 843-846.

Baev, N., Schultze, M., Barlier, I., Cam Ha, D., Kondorosi, E. and Kondorosi, A. (1992) *Rhizobium meliloti nodM* and *nodN* genes are common *nod* genes: *nodM* encodes functions for efficiency of Nod signal production and bacteroid maturation. *J. Bacteriol.* **174**, 7555-7565.

Schultze, M., Quiclet-Sire, B., Kondorosi, É., Virelizier, H., Glushka, J.N., Endre, G., Géro, S.D. and Kondorosi, . (1992) *Rhizobium meliloti* produces a family of sulfated lipo-oligosaccharides exhibiting different degrees of plant host specificity. *Proc. Natl. Acad. Sci. USA* **89**, 192-196.

Soto, M.J., Zorzano, A., Olivares, J. and Toro, N. (1992) Sequence of IS*Rm4* from *Rhizobium meliloti* Strain GR4. *Gene* **120**, 125-126.

Soto, M.J., Zorzano, A., Olivares, J. and Toro, N. (1992) Nucleotide sequence of *Rhizobium meliloti* GR4 insertion sequence ISRm3 linked to the nodulation competitiveness locus nfe. *Plant Mol. Biol.* **20**, 307-309.

Allison, L.A., Kiss, G.B., Bauer, P., Poiret, M., Pierre, M., Savouré, A., Kondorosi, E. and Kondorosi, A. (1993) Identification of two alfalfa early nodulin genes with homology to members of the pea *Enod12* gene family. *Plant Mol. Biol.* **21**, 375-380.

Dusha, I. and Kondorosi, A. (1993) Genes at different regulatory levels are required for the ammonia control of nodulation in *Rhizobium meliloti. Mol. Gen. Genet.* **240**, 435-444.

Esnault, R., Buffard, D., Breda, C., Sallaud, C., El Turk, J. and Kondorosi, A. (1993) Pathological and molecular characterizations of alfalfa interactions with compatible and incompatible bacteria, *Xanthomonas campestris pv. alfalfae and Pseudomonas syringae pv. pisi. Mol. Plant-Microbe Interact.* **6**, 655-664.

John, M., Röhrig, H., Schmidt, J., Wieneke, U. and Schell, J. (1993) *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 625-629.
Mercado-Blanco, J. and Olivares, J. (1993) Stability and transmissibility of the cryptic plasmids of *Rhizobium meliloti* GR4. Their possible use in the construction of cloning vectors for rhizobia. *Arch. Microbiol.* **160**, 477-485.
Schmidt, J., Röhrig, H., John, M., Wieneke, U., Stacey, G., Koncz, C. and Schell, J. (1993) Alteration of plant growth and development by *Rhizobium nodA* and *nodB* genes involved in the synthesis of oligosaccharide signal molecules. *Plant J.* **4**, 651-658.
Soto, N.J., Zorzano, A., Mercado-Blanco, J., Lepek, V., Olivares, J., Toro, N. 1993. Nucleotide sequence and characterization of *Rhizobium meliloti* nodulation competitiveness genes *nfe. J. Mol. Biol.* **229**, 570-576.
Bauer, P., Crespi, M., Szécsi, J., Allison, L.A., Schultze, M., Ratet, P., Kondorosi, E. and Kondorosi, A. (1994) Alfalfa *Enod12* genes are differentially regulated during nodule development by Nod factors and *Rhizobium* invasion. *Plant Physiol.* **105**, 585-592.
Blondon, F., Marie, D., Brown, S. and Kondorosi, A. (1994) Genome size and base composition in *Medicago sativa L.* and *Medicago truncatula* Gaertn. species. *Genome*, (in press).
Cren, M., Kondorosi, A. and Kondorosi, E. (1994) An insertional point mutation inactivates NolR repressor in *Rhizobium meliloti* strain 1021. *J. Bacteriol.*, **176**, 518-519.
Cren, M., Kondorosi, A. and Kondorosi, E. (1994) NolR controls expression of the *Rhizobium meliloti* nodulation genes involved in the core Nod factor synthesis. *Mol. Microbiol.*, under revision.
Crespi, M.D., Jurkevitch, E., Poiret, M., d'Aubenton-Carafa, Y., Petrovics, G., Kondorosi, E. and Kondorosi, A. (1994) *Enod40*, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. *EMBO J.*, (in press).
Csanádi, G., Szécsi, J., Kaló, P., Kiss, P., Endre, G., Kondorosi, A., Kondorosi, É and Kiss, G.B. (1994) *Enod12*, an early nodulin gene, is not required for nodule formation and efficient nitrogen fixation in alfalfa. *Plant Cell*, **6**, 201-213.
Mercado-Blanco, J. and Olivares, J. (1994) A protein involved in stabilization of a large non-symbiotic plasmid of *Rhizobium meliloti* shows homology to eukaryotic cytoskeletal proteins and DNA-binding proteins. *Gene*, **139**, 133-134.
Mercado-Blanco, J. and Olivares, J. (1994) The large non-symbiotic plasmid pRmeGR4a of *Rhizobium meliloti* GR4 encodes a protein involved in replication that has homology with the RepC protein of *Agrobacterium* plasmids. *Plasmid*, (in press).
Savouré, A., Magyar, Z., Pierre, M., Brown, S., Schultze, M., Dudits, D., Kondorosi, A. and Kondorosi, E. (1994) Activation of the cell cycle machinery and the isoflavonoid biosynthesis pathway by active *Rhizobium meliloti* Nod signal molecules in *Medicago* microcallus suspensions. *EMBO J.*, **13**, 1093-1102.
Savouré, A., Kaló, P., Fehér, A., Petrovics, G., Csanádi, G., Kiss, G.B., Brown, S., Kondorosi, A. and Kondorosi, E. (1994) Identification, characterization and mapping of a full length cyclin gene *CycIIIMs* in *Medicago sativa. Plant Mol. Biol.*, (submitted).
Schultze, M., Kondorosi, E., Ratet, P., Buiré, M. and Kondorosi, A. (1994) Cell and molecular biology of *Rhizobium*-plant interactions. *Int. Rev. Cytol.*, (in press).
Soto, M.J., van Dillewijn, P., Olivares, J. and Toro N. (1994) Ornithine cyclodeaminase activity in *Rhizobium meliloti. FEMS Microbiol. Lett.*, (in press).
Soto, M.J., Zorzano, A., Mercado-Blanco, J., López-Lara, I.M., Olivares, J. and Toro, N. (1994) Identification of a novel *Rhizobium meliloti* nodulation efficiency *nfe* gene. *Mol. Plant-Microbe Interact.*, under revision.
Staehelin, C., Schultze, M., Kondorosi, É., Mellor, R.B., Boller, T. and Kondorosi, A. (1994) Structural modifications in *Rhizobium meliloti* Nod factors influence their stability against hydrolysis by root chitinases. *Plant J.*, **5**, 319-330.

The molecular basis of pathogenicity, avirulence and resistance in the interaction between the fungal pathogen *Cladosporium fulvum* (syn. *Fulvia fulva*) and tomato (BIOT CT-900163)

COORDINATOR:
A. P.J.G.M. de WIT, Wageningen Agricultural Univ., Wageningen, NL
PARTICIPANTS
B. R.P. OLIVER, University of East Anglia, Norwich, GB
C. H.W.J. van den BROEK, Wageningen Agricult. Univ., Wageningen NL
D. L.S. MELCHERS, MOGEN Int. N.V., Leiden, NL
E. S.C. FRY, University of Edinburgh, Edinburgh, GB
F. F. CERVONE, Università 'La Sapienza', Rome, I

BACKGROUND INFORMATION

The molecular basis underlying the pathogenicity of the biotrophic fungus *Cladosporium fulvum* is still obscure. It is a highly specialized pathogen, limited to infect only a few species in the *Lycopersicon genus*. Studies with pathogenic bacteria and fungi have shown that several genes are activated during penetration of the host and during pathogenesis. The isolation of *in planta* induced genes of *C. fulvum*, which likely play an important role in the infection process and therefore could be considered as putative pathogenicity factors, is one of the topics of our research. Host resistance is assumed to be based on recognition of the pathogen, resulting from direct interaction between avirulence gene *(Avr)* products (race-specific elicitors) and resistance gene *(Cf)* products (host-specific receptors). Upon recognition, a hypersensitive-response (HR) takes place in the host and fungal development is arrested. We are interested in the isolation of avirulence gene products and their encoding genes, as this establishes a basis for studies on signalling events between host and pathogen.

Besides its direct effect on *C. fulvum*, HR induces (systemic) resistance against a range of other pathogens. The phenomenon of induced resistance in plants is accompanied by the induction of the synthesis of a large number of pathogenesis-related (PR) proteins, including proteins with the ability to inhibit growth of fungi *in vitro*. In literature it is proposed that these proteins are involved in defense against fungal infections *in vivo*.

HR is probably not only the result of specific recognition of proteinaceous fungal elicitors; it can also be induced by plant and fungal cell wall components. A protein which inhibits fungal endo-polygalacturonases (PGIP) is present in the cell wall of all dicotyledonous plants that have been examined. Interaction between endopolygalacturonase and PGIP *in vitro* results in the accumulation of oligogalacturonides active as elicitors of plant defense responses. It has been suggested that PGIP is invoved in plant resistance to fungal pathogens.

OBJECTIVES AND PRIMARY APPROACHES

Lab A focused attention on two main objectives; unraveling of the different components of the defense response of tomato plants resistant to *C. fulvum* and isolation and characterisation of avirulence genes of *C. fulvum*.

Lab B concentrated on testing of approaches for the isolation of genes controlling the pathogenicity of *C. fulvum*. In order to investigate the biological function of putative pathogenicity genes of *C. fulvum*, lab C adopted gene disruption and gene replacement methodologies to study the function of these genes for *C. fulvum*.

The objective of the fungal resistance project in lab D is the development of strategies for the control of fungal diseases in transgenic crop plants. One strategy consists of the identification of antifungal proteins and the constitutive expression of the corresponding genes in transgenic plants.

The work in lab E was aimed at defining the effects of hitherto unstudied cell wall fragments on plant metabolism. The lab also undertook some investigations into the structural features of the novel oligosaccharins.

The aim of the research carried out by lab F was to elucidate the role of PGIP in plant-fungus interactions with a view to using the gene encoding PGIP for plant transformation to improve resistance to phytopathogenic fungi. To this purpose, the *pgip* gene was cloned and transferred to tomato plants to obtain transgenic plants expressing high levels of PGIP.

RESULTS AND DISCUSSION

Lab A. From apoplastic fluid (AF), isolated from a compatible *C. fulvum*-tomato interaction, a 26kD and 27kD acidic chitinase and a 35kD acidic 1,3-β-glucanase were purified, while from homogenates of infected leaves two intracellular, basic chitinases (30kD and 32kD) and one basic, 33kD, 1,3-β-glucanase were obtained. Assessment of the antifungal activity of the purified chitinases and 1,3-β-glucanases revealed that the test fungus, *Trichoderma viride* was highly sensitive to the intracellular, basic isoforms. Germinated conidia of *C. fulvum*, however, were found not to be inhibited by the hydrolytic enzymes. If growth of *C. fulvum* is indeed not affected by the chitinases and 1,3-β-glucanases, this would suggest that HR itself, including host cell death, callose deposition, oxidative burst and other early defense responses, is sufficient for the arrest of fungal growth.

Avirulence gene *Avr*4 of *C. fulvum* was shown to encode a pre-pro-protein of 135 amino acids, containing a signal peptide for extracellular targeting and a stretch of amino acids which is cleaved off by plant and/or fungal proteases. The resulting elicitor protein of 106 amino acids contains 8 cysteine residues. In races of *C. fulvum* virulent on tomato genotype Cf4, a gene highly similar to the *Avr*4 gene is present, except that, as a result of a point mutation, one cysteine residue (codon TGT) has been replaced by a tyrosine residue (codon TAT). In one isolate virulent on tomato genotype Cf4, a deletion of one base-pair in the 5'-part of the open reading frame encoding the mature AVR4 protein, resulted in a frame shift. These results show that *Cf4*-specific resistance is circumvented by a single base-pair change in avirulence gene *Avr*4 of *C. fulvum*.

Lab B. To facilitate the mapping of mutated genes, a method of genetic analysis based on an induced parasexual cycle has been developed. This involves the fusion of protoplasts followed by spontaneous rehaploidisation. Markers are scored in the parasexual progeny. Eleven linkage groups have been identified. Telomeric DNA has been cloned and characterised. Molecular evidence for a high level of recombination has been obtained. Transformation may provide a means to simultaneously inactivate and tag pathogenicity genes. Transformation with pAN8-1 and pAN7-1 leads to integration of the plasmid into the *C. fulvum* genome with apparent randomness; the pathogenicity phenotype of nearly 1000 transformants has been assessed. Whilst mutations in extracellular enzymes appear common,

mutations clearly affecting pathogenicity have proved elusive. The identification of fungal genes that are induced during growth *in planta* would greatly enhance our knowledge of the molecular and physiological basis of pathogenicity. Five *in planta*-induced clones were isolated. Two of the clones are alcohol dehydrogenase and aldehyde dehydrogenase.

Work in Wageningen and Leiden has indicated that *C. fulvum* is resistant to the antifungal chitinases and 1,3-β-glucanases produced in infected plants. Work in Norwich and during a visit to Wageningen suggests that *C. fulvum* excretes an inhibitor, which has been partially purified by reversed-phase chromatography.

Work in Wageningen had indicated that photosynthetic sucrose was converted during the infection into glucose and fructose. Mannitol dehydrogenase, the enzyme which converts fructose into mannitol, was purified from *C. fulvum*. The equilibrium constant, Km values and substrate specificity were measured. These results are consistent with the notion that the fungus converts 'spare' carbohydrate into mannitol for use as a storage compound. This hypothesis will be tested by cloning the gene and creating disruption mutants.

Lab C. Races of *C. fulvum*, avirulent on tomato genotypes carrying the resistance gene *Cf9*, were used to develop the *avr9* disruption procedure. Selected uridine auxotrophic mutants of these strains were transformed with a plasmid containing the *avr9* genomic region in which the open reading frame was replaced by the *pyrG* gene from *Aspergillus nidulans*. Transformants in which the entire *avr9* coding sequence was deleted as a result of a gene replacement event were shown to be able to successfully infect *Cf9* tomato genotypes. (In previous studies it was shown that integration of the *avr9* gene into the genome of a race that did not contain the gene, changed this race from virulent to avirulent on tomato genotype *Cf9*). These results confirmed that the cloned *avr9* sequence is the only genetic factor that is responsible for the inability of certain *C. fulvum* races to infect *Cf9* tomato plants and that the AVR9 peptide is the only compound from intercellular fluids that induces necrosis in *Cf9* plants.

The gene replacement strategy has also been followed to knock out genes encoding putative extracellular pathogenicity factors (*ecp* genes) of *C. fulvum*. The *ecp*2 gene was completely deleted from the genome and replaced by a phleomycin resistance cassette; the *ecp*1 gene was disrupted by deletion of the *ecp*1 promoter and flanking sequence and replaced by a hygromycin B resistance gene cassette. Both types of disruption mutants were still pathogenic on tomato seedlings. This clearly indicates that these *C. fulvum* genes are not essential for pathogenicity on tomato. To see the effect of disruption of both the *ecp*1 gene and the *ecp*2 gene on pathogenicity, fusions between the *ecp*1 and *ecp*2 mutants will be analysed to study the possible involvement of cooperatively working ECPs in pathogenicity of *C. fulvum*.

Lab D. MOGEN has developed the expertise to transform and to regenerate several crop species, among which tomato, and has constructed expression and transformation vectors. The tomato transformation technology was improved significantly in the contract period and was used for the introduction of a chitinase/glucanase multiple gene construct and constructs provided by participants. Chimeric gene constructs were made containing the tobacco chitinase and 1,3-β-glucanase genes under the control of the constitutive *cauliflower mosaic virus* (CaMV) 35S promoter and were introduced into tomato using an *Agrobacterium*-mediated transformation procedure. Transgenic plants were assayed for resistance to the phytopathogen *Fusarium oxysporum* f.sp. *lycopersici* race 1, a soil-borne

fungus that infects the vascular tissue of tomato after penetration of roots through wound sites. We demonstrated fungal resistance in tomato by the simultaneous expression of a class I chitinase and a class I 1,3-β-glucanase gene from tobacco. Transgenic plants expressing either of these genes alone were not protected. These results are consistent with the observation that class I chitinases and class I 1,3-β-glucanases synergistically inhibit growth of fungi *in vitro* and provide the first experimental support to the hypothesis that such synergism can contribute to enhanced fungal resistance *in vivo*.

Lab E. Two polysaccharides isolated from tomato cell walls, rhamnogalacturonan-II (RG-II) and glucuronoxylan (GX) were found to have strong inhibitory effects on [^{14}C]leucine uptake of a tomato cell culture. RG-II appeared to affect the uptake of some amino acids (besides [^{14}C]leucine also [^{14}C]glutamate) but not all (e.g. not [^{14}C]phenylalanine), indicating interference with the mechanism of action of specific permease systems. The most active material in the RG-II preparations was analysed by HPLC, after complete acid hydrolysis and was shown to contain, amongst others, the unusual sugar residues, characteristic of RG-II, apiose, 2-O-methylxylose, 2-O-methylfucose and ketodeoxyoctulosonic acid (KDO). In order to check for the possible occurrence of enzyme activities, produced by *C. fulvum*, with the ability to generate oligosaccharins from the host plant cell wall, we tested the effect of culture filtrate of *C. fulvum* on radioactively labelled plant cell walls. The data indicated the production of small amounts of specific oligosaccharides from the plant cell walls. These have been partially characterised and do not include material characteristic of RG-II. Conversely, we tested the ability of apoplastic fluid from tomato leaves to solubilise oligosaccharides from ^{14}C-labelled hyphal walls of *C. fulvum*. This again indicated the presence of oligosaccharide products, but it was not possible to relate their structures to any oligosaccharins identified by the work described above.

Lab F. The *pgip* gene of *Phaseolus vulgaris* has been cloned, characterized and shown to encode a 342-amino acid polypeptide including a signal peptide for extracellular targeting. The PGIP amino acid sequence exhibits significant internal homology in a domain spanning from position 69 to 326. This domain exhibits a modular structure and can be divided into a set of 10.5 tandemly repeated units, each derived from modifications of a 24-amino acid leucine-rich peptide. A chimeric gene was constructed with the coding region of the bean *pgip* gene, under control of the CaMV 35S promoter and introduced into tomato plants by *A. tumefaciens*-mediated transformation. Transgenic tomato plants expressing high levels of bean PGIP were obtained and are currently being tested for increased resistance to various fungi. Transgenic tobacco plants expressing *pgip*-promoter/GUS gene fusions have also been obtained. Accumulation of *pgip* transcripts was induced in suspension-cultured bean cells, following addition of elicitor-active oligogalacturonides and fungal glucan. Accumulation of *pgip* mRNA also occurred in bean hypocotyls in response to wounding or treatment with salicylic acid. Moreover, when different bean cultivars are infected with different races of *Colletotrichum lindemuthianum*, a rapid accumulation of *pgip* mRNA at the site of infection correlates with the appearance of the hypersensitive response in incompatible interactions, while a more delayed increase, coincident with the onset of lesion formation, occurs in the compatible interaction.

MAJOR SCIENTIFIC BREAKTHROUGHS

(A) The second fungal avirulence gene, *Avr*4, of *C. fulvum* has been cloned. A single basepair change in this gene changes the fungus from avirulent to virulent on tomato plants carrying resistance gene *Cf4*.
(B) Eleven linkage groups have been identified in *C. fulvum*. Telomeric DNA of *C. fulvum* has been cloned and characterized and telomere-linked RFLPs have been shown to map to linkage group ends.
(C) Avirulence gene *avr*9 has been successfully disrupted in avirulent wild-type races; transformants became virulent on tomato plants carrying resistance gene *Cf9*. Disruption of the *ecp*1 or the *ecp*2 gene (genes encoding extracellular proteins, that are specifically induced *in planta*) revealed that these genes cannot be considered to encode strong pathogenicity factors for *C. fulvum*.
(D) In tomato, resistance to *F. oxysporum* can be obtained by the simultaneous expression of a class I chitinase and a class I 1,3-β-glucanase gene from tobacco. It was demonstrated that overexpression of antifungal proteins is a feasible approach for enhancing fungal resistance in economically important crop plants.
(E) A sub-fraction of rhamnogalacturonan-II has been shown to possess biological activity on tomato cells. Healthy tomato leaf cells produce enzymes that liberate oligosaccharides from *C. fulvum* cell walls and enzymes produced by *C. fulvum* generate specific oligosaccharides from tomato cell walls.
(F) The first *pgip* gene has been cloned from bean and transferred to tomato plants. Expression of the gene has been shown to be up-regulated in response to elicitors, wounding and infection. Rapid accumulation of *pgip* mRNA at the site of infection has been shown to correlate with the appearance of the hypersensitive response in incompatible interactions. PGIP is the first plant leucine-rich protein shown to participate in plant-fungus communication.

MAJOR COOPERATIVE LINKS

Intensive exchange of expertise and material was realised between the two Wageningen labs; after his PhD study in lab A, Van den Ackerveken spent a postdoc period of one year in lab C. Richard Laugé is appointed on a combined project of the two groups. Strains and plasmid constructs were also exchanged with lab B. R. Oliver spent 2 weeks in Wageningen. Expression vectors, binary plasmids and *Agrobacterium* strains were supplied by lab D to A and F. Lab D produced, using *Agrobacterium*-mediated transformation, a large number of transgenic tomato and or tobacco plants constitutively expressing either *avr*9, *pgip* or chitinase/glucanase genes. A PhD-student from group F spent 60 days in group D for the expression analysis of PGIP transgenic tomato plants. These transgenic plants were tested for increased fungal resistance by group D. Labs A and B supplied culture filtrates of *C. fulvum* to lab E. Lab A also supplied this lab with ^{14}C-labelled hyphal wall material and samples of apoplastic fluid from healthy and *C. fulvum*-infected tomato leaves.

PUBLICATIONS

Joint publications

De Wit, P.J.G.M., Van den Ackerveken, G.F.J.M., Vossen J.P.M.J., Joosten, M.H.A.J., Cozijnsen, T.J.,Honee, G., Wubben, J.P., Danhash, N., Van Kan, J.A.L., Marmeisse, R. & Van den Broek, H.W.J. (1993) Molecular cloning and functions of avirulence and pathogenicity genes of the tomato pathogen *Cladosporium fulvum*. In: Nester, E.W. & Verma, D.P.S., (eds), Advances in molecular genetics of plant-microbe interactions, Vol. 2, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993, pp. 289-298.

De Wit, P.J.G.M., Van Den Ackerveken, G.F.J.M., Vossen, P.M.J., Joosten M.H.A.J., Cozijnsen, A.J., Honee, G., Wubben, J.P., Danhash, N., Van Kan, J.A.L., Marmeisse, R. and Van Den Broek, H.W.J. (1993). Avirulence genes of the tomato pathogen *Cladosporium fulvum* and their exploitation in molecular breeding for disease-resistant plants. In: Developments in Plant Pathology; Mechanisms of plant defense responses (eds. B. Fritig and M. Legrand), Kluwer Academic Publishers, Dordrecht/Boston/London, pp. 24-32.
De Wit, P.J.G.M., Joosten, M.H.A.J., Honée, G., Wubben, J.P., Van den Ackerveken, G.F.J.M. & Van den Broek, H.W.J. (1994). Molecular communication between host plant and the fungal tomato pathogen *Cladosporium fulvum*. Anthonie van Leeuwenhoek J. of Microbiology, (in press).
Marmeisse, R., Van den Ackerveken, G.F.J.M., Goosen, T., De Wit, P.J.G.M. & Van den Broek, H.W.J. (1993). Disruption of the avirulence gene *avr9* in two races of the tomato pathogen *Cladosporium fulvum* causes virulence on tomato genotypes with the complementary resistance gene *Cf9*. Mol. Plant-Microbe Interact., **6**, 412-417.
Marmeisse, R., Van den Ackerveken, G.F.J.M., Goosen, T., De Wit, P.J.G.M., & Van den Broek, H.W.J. (1994). The *in planta* induced *ecp*2 gene of the tomato pathogen *Cladosporium fulvum* is not essential for pathogenicity. Curr. Genet., **26**, 245-250.
Van den Ackerveken, G.F.J.M., Dunn, R.M., Cozijnsen, T.J., Vossen, J.P.M.J., Van den Broek, H.W.J. & De Wit, P.J.G.M. (1994). Nitrogen limitation induces expression of the avirulence gene *avr9* in the tomato pathogen *Cladosporium fulvum*. Mol. Gen. Genet. **243**, 277-285.
Van den Ackerveken, G.F.J.M., Marmeisse, R., Laugé, R., Goosen, T. De Wit, P.J.G.M. & Van den Broek, H.W.J. (1994). Disruption of *ecp*1 in *Cladosporium fulvum*: Do the *ecp* genes have complementary functions? (in preparation)

Individual publications

Aldington, S. & Fry, S.C. (1992). Plant cell wall-lysing enzymes in the apoplast of tomato leaves infected with *Fulvia fulva*. Can. J. Bot., **70**, 607-613.
Aldington, S. & Fry, S.C. (1993). Oligosaccharins. Adv. Bot. Res., **19**, 1-101.
Aldington, S. & Fry, S.C. (1994). Rhamnogalacturonan-II — a biologically active fragment. Journal of Experimental Botany, **45**: 287-293.
Aldington, S. & Fry, S.C. (1994). Effects of diverse polysaccharides and oligosaccharides on [^{14}C]leucine incorporation by tomato cell suspension cultures. J. Plant Physiol., (in press).
Arnau, J.A. & Oliver, R.P. (1993). Inheritance and alteration of transforming DNA during an induced parasexual cycle in the imperfect fungus *Cladosporium fulvum*. Curr. Genet., **23**, 508-511.
Arnau, J.A. & Oliver, R.P. (1994). Use of RAPD markers in the genetic analysis of the plant pathogenic fungus, *Cladosporium fulvum*. Curr. Genet., **25**, 438-444.
Arnau, J.A. & Oliver, R.P. (1994). Genetic analysis in the imperfect fungus *Cladosporium fulvum*; the development of a map-based strategy for the cloning of genes involved in pathogenicity. In: Bos C.J. (ed) Fungal genetics, principles and practice. (in press)
Bellincampi, D., Salvi, G., De Lorenzo, G., Cervone, F., Marfa', V., Eberhard, S., Darvill, A. & Albersheim, P. (1993). Oligogalacturonides inhibit the formation of roots on tobacco explants. Plant J., **4**, 207-213.
Bergmann, C., Ito, Y., Singer, D., Albersheim, P., Darvill, A.G., Benhamou, N., Nuss, L., Salvi, G., Cervone, F. & De Lorenzo G. (1994). Polygalacturonase-inhibiting protein accumulates in *Phaseolus vulgaris L.* in response to wounding, elicitors and fungal infection. Plant J., (in press).
Caprari, C., Richter, A., Bergmann, C., Lo Cicero, S., Salvi, G., Cervone, F. & De Lorenzo, G. (1993). Cloning and characterization of the gene encoding the endopolygalacturonase of *Fusarium moniliforme*. Mycol. Res. **97**, 497-505.
Caprari, C., Bergmann, C., Migheli, Q., Salvi, G., Albersheim, P., Darvill, A., Cervone, F. & De Lorenzo, G. (1994). *Fusarium moniliforme* secretes four endopolygalacturonases derived from a single gene product. Physiol. Mol. Plant Pathol., (in press).
Cervone, F., De Lorenzo, G., Caprari, C., Clark, A.J., Desiderio, A., Devoto, A., Leckie, F., Nuss, L.,Salvi, G. & Toubart, P. (1993). The interaction between fungal endopolygalacturonase and plant cell wall PGIP (Polygalacturonase-Inhibiting Protein). In: Fritig, B. & Legrand, M. (eds), mechanisms of plant defense response, pp. 64-67.

Clark, L., Nuss, L., Mahi, A., Grisvard, J., Dron, M., Cervone, F. & De Lorenzo, (1994). Polygalacturonase-inhibiting protein (PGIP) is rapidly and spatially up-regulated during incompatible interactions of *Phaseolus vulgaris L.* and *Colletotrichum lindemuthianum.* (submitted).
Coleman, M.J., McHale, M.T., Arnau, J., Watson, A. & Oliver, R.P. (1993). Cloning and characterisation of telomeric DNA from *Cladosporium fulvum*. Gene, **132**, 67-73
Coleman, M.J., Faber, B. & Oliver, R.P. (1994/5). Many *Cladosporium fulvum* genes induced by growthin the host tomato are also induced by starvation *in vitro*. (in preparation).
Coleman, M.J., & Oliver, R.P. (1994/5). Expression of *Cladosporium fulvum* alcohol dehydrogenase andaldehyde dehydrogenase genes during growth in the host tomato. (in preparation).
Cornelissen, B.J.C. & Melchers, L.S. (1993). Strategies for control of fungal diseases with transgenic plants. Plant Physiol., **101**, 709-712.
Darvill, A. Bergmann, C., Cervone, F., De Lorenzo, G., Ham, K.S., Spiro, M.D., York, W.S. & Albersheim, P. (1994). Oligosaccharins involved in plant growth and host/pathogen interactions. In: Bowles, D.J. (ed) Molecular botany: signals and the environment. Portland Press, London, (in press).
De Lorenzo, G., Cervone, F., Bellincampi, D., Caprari, C., Clark, A.J., Desiderio, A., Devoto, A., Forrest, R., Leckie, F., Nuss, L. & Salvi, G. (1994). Polygalacturonase, PGIP and oligogalacturonides in cell-cell communication. Biochem. Soc. Transact., (in press).
De Lorenzo, G., Castoria, R., Bellincampi, D. & Cervone, F. (1994/5). Fungal invasion enzymes and their inhibition. In: Carroll, G. & Tudzynski, P. (eds), The Mycota, Vol 6, Plant relationships. Springer-Verlag, Berlin, (in press).
De Wit, P.J.G.M. (1992). Molecular characterization of gene-for-gene systems in plant-fungus interactions and the application of avirulence genes in control of plant pathogens. Ann. Rev. Phytopathol., **30**, 391-418.
Frediani, M., Cremonini, R., Salvi, G., Caprari, C., Desiderio, A., D'ovidio, R., Cervone, F. & De Lorenzo, G.(1993). Cytological localization of the *pgip* genes in the embryo suspensor cells of *Phaseolus vulgaris L.* Theor. Appl. Genet., **87**, 369-373.
Fry, S.C., Aldington, S., Hetherington, P.R. & Aitken, J. (1993). Oligosaccharides as signals and substrates in the plant cell wall. Plant Physiol., **103**: 1-5.
Joosten, M.H.A.J., Cozijnsen, A.J. & De Wit, P.J.G.M. (1994). Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. Nature, **367**, 384-386
Joosten, M.H.A.J., Verbakel, H.M., Nettekoven, M.E., Van Leeuwen, J, Van der Vossen, R.T.M. & De Wit, P.J.G.M. (1994). The fungal pathogen *Cladosporium fulvum* is not sensitive to the chitinase and 1,3-β-glucanase defence proteins from its host, tomato. Physiol. Mol. Plant Pathol., **45**, (in press).
Joosten, M.H.A.J., Honée, G., Van Kan, J.A.L. & De Wit, P.J.G.M. (1994/5). The gene-for-gene concept in plant-pathogen interactions: tomato-*Cladosporium fulvum*. In: Carroll, G. & Tudzynski, P. (eds), The Mycota, Vol 6, Plant relationships. Springer-Verlag, Berlin, (in press).
Leckie, F., Devoto, A., Cervone, F. & De Lorenzo, G. (1994). The promoter of the gene encoding the polygalacturonase-inhibiting protein (PGIP). (submitted).
Melchers, L.S., Apotheker-de Groot, M., Van der Knaap, J.A., Ponstein, A.S., Sela-Buurlage, M.B., Cornelissen, B.J.C., van den Elzen, P.J.M. & Linthorst, H.J.M. (1994). A new class of tobacco chitinases homologous to bacterial exo-chitinases displays antifungal activity. Plant J., **5**, 469-480.
Melchers, L.S., Sela-Buurlage, M.B., Vloemans, S.A., Woloshuk, C.P., Van Roekel, J.S.C., Pen. J., Van den Elzen, P.J.M. and Cornelissen, B.J.C. (1993). Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and β-1,3-glucanase in transgenic plants. Plant Mol. Biol., **21**, 583-593.
Melchers, L.S., Ponstein, A.S., Sela-Buurlage, M.B., Vloemans, S.A. and Cornelissen. B.J.C. (1993). *In vitro* anti-microbial activities of defense proteins and biotechnology. In B. Fritig, M. Legrand (eds), Mechanisms of Plant defense Responses, vol. 2, Kluwer Academic Publishers, pp. 401-410.
Noldner, P., Coleman, M.J., Faulks, R. & Oliver, R.P., (1994/5). Purification and characterization of the mannitol dehydrogenase of the fungal tomato pathogen, *Cladosporium fulvum*. (in preparation.

Oliver, R.P. (1993). The use of nucleic acid-based methods in the detection and identification of plant pathogens. In: Fox, R.T.V. (ed.) Principles of diagnostic methods in plant pathology. CAB International.
Oliver, R.P., Farman, M.L., Jones, J.D.G. & Hammond-Kosack, K.E. (1993). Use of fungal transformants expressing β-glucuronidase activity to detect and measure hyphal biomass in infected plant tissues. Mol Plant-Microbe Int., **6**, 521-525.
Oliver, R.P. (1994/5). Use of AFLPs to examine the evolution of isolates of *Cladosporium fulvum*. (iin preparation).
Ponstein, A.S., Bres-Vloemans, S.A., Sela-Buurlage, M.B., van den Elzen, P.J.M., Melchers, L.S. & Cornelissen, B.J.C.(1994). A novel pathogen- and wound-inducible tobacco protein with antifungal activity. Plant Physiol., **104**, 109-118.
Redgwell, R.J. & Fry, S.C. (1993). Xyloglucan endotransglycosylase activity increases during kiwifruit *(Actinidia deliciosa)* ripening: implications for fruit softening. Plant Physiol., **103**: 1399-1406.
Sela-Buurlage, M.B., Ponstein, A.S., Vloemans, S.A., Melchers, L.S., Van den Elzen, P.J.M. & Cornelissen, B.J.C. (1993). Only specific chitinases and β-1,3-glucanases exhibit antifungal activity. Plant Physiol., **101**, 857-863.
Van den Ackerveken, G.F.J.M., Van Kan, J.A.L., Joosten, M.H.A.J., Muisers, J.M., Verbakel, H.M. and De Wit, P.J.G.M. (1993). Characterization of two putative pathogenicity genes of the fungal tomato pathogen *Cladosporium fulvum*. Mol. Plant-Microbe Int., **6**, 210-215.
Van den Elzen, P.J.M., Jongedijk, E., Melchers, L.S., & Cornelissen, B.J.C. (1993). Virus and Fungal Resistance: From laboratory to field. Philosophical Transactions of the Royal Society of London Series B — Biological Sciences **342**, 271-278.
Van Kan, J.A.L., Joosten, M.H.A.J., Wagemakers, C.A.M., Van den Berg-Velthuis, G.C.M. & De Wit, P.J.G.M. (1992). Differential accumulation of mRNAs encoding extracellular and intracellular PR proteins in tomato induced by virulent and avirulent races of *Cladosporium fulvum*. Plant Mol. Biol., 20: 513-527
Van Kan, J.A.L., Joosten, M.H.A.J., Van den Ackerveken, G.F.J.M. & De Wit, P.J.G.M. (1994). Molecular characterization of avirulence determinants of the tomato pathogen *Cladosporium fulvum*. In: Kohmoto, K. & Yoder, O.C. (eds), Host-specific toxin biosynthesis, receptor and molecular biology. Sogo Printing and Publishing Co., Ltd., Tottori, Japan. pp. 251-261.

Genetic and molecular approaches to the physiology of bacteroids in relation to the plant nodule metabolism (BIOT CT-900166)

COORDINATOR:
P. BOISTARD, INRA-CNRS, Toulouse, F
PARTICIPANTS:
M. IACCARINO, IIGB, CNR, Naples, I
R. KRON-MORELLI, Agrifutur, Brescia, I
M. O'CONNELL, Dublin City University, IRL
F. O'GARA, University College, Cork, IRL
A. PÜHLER and U. PRIEFER, University of Bielefeld and RWTH Aachen, D

BACKGROUND INFORMATION

Symbiotic nitrogen (N_2) fixation is an economically and environmentally acceptable alternative to the use of fertilizers, worldwide. Its improvement ultimately depends on a better understanding of the regulatory circuits operating in both the legume and the *Rhizobium* partner and of their mutual interactions.

OBJECTIVES AND PRIMARY APPROACHES

Our project aims at an overall picture of the regulatory networks operative in the nodule that control nitrogen fixation by *Rhizobium* bacteroids. We concentrate on the interconnected roles of oxygen limitation, carbon and nitrogen supply on nitrogen fixation gene expression and activity.

RESULTS AND DISCUSSION

1. Oxygen regulation of nitrogen fixation genes (Aachen, Bielefeld, Toulouse)

In *R. meliloti* where the complete regulatory cascade was identified genetically, in vitro assays have been developed in order to study the mechanism by which oxygen concentration affects transcription of nitrogen fixation genes. We have found that the FixL sensor protein dramatically enhances FixJ transcriptional activity by phosphorylation in anaerobic conditions (15). We have also shown that a primary effect of FixJ phosphorylation is to enhance its binding at two specific sites at the *fixK* promoter (26). Finally, we have found that, unexpectedly, a mutant FixJ protein in which the normal phosphorylation site has been mutated, is activated by an alternative phosphorylation pathway involving the FixL protein but independently of the oxygen status (20). The physiological significance of this alternate activation pathway of FixJ is currently being evaluated.

The knowledge about oxygen regulatory components involved in the regulation of nitrogen fixation genes in *R. leguminosarum* bv. *viciae* was very poor at the beginning of the project. The only oxygen regulatory protein known at that time was NifA which, like in all other nitrogen fixing bacteria, activates a number of *nif* and *fix* genes (especially those necessary for an active nitrogenase complex). We have now identified a number of genes that, in addition to *nifA*, might play a role in the oxygen regulatory cascade in *R.l.* bv. *viciae*. *fnrN* is a *R. meliloti fixK*-like gene which is preferentially expressed under microaerobic conditions (1). However, in contrast to *R. meliloti fixK*, microaerobic expression of fnrN is independent of the *R. meliloti fixLJ* operon. Instead it is subject to positive autoregulation. A further difference with *R. meliloti fixK* is that *fnrN* is not only regulated by oxygen at the transcriptional level but also at the level of protein activity (10,16). This difference

can be attributed to the presence of an N-terminal cysteine cluster which is missing in *R. meliloti* FixK. In addition to *fnrN*, which is located on the chromosome, we found a second *fnrN/fixK* like gene in *R. leguminosarum*, which is located on the non-nodulation plasmid pRlVF39c upstream of the *fixNOQP* operon. This FixK homologue shows about 32% similarity to *Rm*FixK and does not contain an N-terminal cysteine motif. Downstream of *Rl*fixK, we identified a gene with conspicuous homology to *RmfixL*. Further sequence analysis of the region downstream of the *fixL*-like gene revealed the presence of a third open reading frame with similarity to *RmfixJ*. Interestingly, this open reading frame, presently named ORF131, consists only of the N-terminal receiver domain of FixJ whereas the C-terminal transcriptional activator domain of FixJ is completely missing. Therefore the putative gene product of *orf131* resembles the regulatory protein SpoOF from *Bacillus subtilis* and might represent an intermediate phosphorylation carrier in the signal transduction pathway. This will be tested by mutational analyses and phosphorylation studies (in collaboration with Toulouse). We introduced several mutations into the *fixL* and *fixK* genes of *R. leguminosarum*. These mutants are currently being tested on *Vicia hirsuta* to test the involvement of these genes in symbiotic nitrogen fixation. Future work is required to determine the interrelation between these components in the oxygen regulatory network leading to the expresssion of nitrogen fixation genes in *R. leguminosarum*. Most intriguingly, none of the so far identified target genes of the *fixLJ* system in *R. meliloti* (i.e. *nifA, fixK*) can be induced in *R.l. viciae* thus suggesting that the regulatory pathways in these two symbionts may differ significantly.

Another aspect of our work has been the identification of two *fixNOQP* operons in *R.l.* bv. *viciae* (16). *fixNOQP* genes have been first identified in *R. meliloti* and, more recently, in *B. japonicum*. Both aminoacid sequence and biochemical analyses indicate that the *fixNOQP* operon most likely encodes an oxidase complex with high affinity for oxygen that would allow bacterial respiration under microoxic conditions, as in the nodule environment. Expression studies have shown that the pSym-located copy of *fixNOQP* is induced under microaerobic conditions as in *R. meliloti*. We could also show that *fixN*-homologous genes are not only present in rhizobia but also in other nitrogen fixing bacteria (i.e. *Rhodobacter capsulatus*) and also in the non-fixing plant-pathogen *Agrobacterium tumefaciens* (29). These data support the hypothesis that the *fixN* gene product is not specific for nitrogen fixation but is involved in bacterial respiration under reduced oxygen availability.

2. Carbon metabolism and nitrogen fixation (Cork and Brescia):

2.1 Alternative symbiotic activator of dctA:

Dicarboxylic acids (dcAs) constitute the major carbon source for the nitrogen fixing bacteroids. Accordingly, mutants in *dctA*, the gene that encodes the membrane-located dcA transporter, are Fix$^-$. Under free-living conditions, the expression of the *R. meliloti dctA* gene is controlled by the divergently transcribed *dctBD* genes. This two-component regulatory system consists of the sensor *DctB* and the transcriptional regulator, DctD. *In vitro* phosphorylation assays indicated that signal transduction between *DctB* and *DctD* takes place by means of phospho-transfer between the two proteins. During symbiosis, *dctA* is expressed in the nodule even in the absence of *dctBD* genes thus suggesting the existence of an alternative symbiotic activator of *dctA*

ntrC is a major regulator of nitrogen metabolism in rhizobia as in enteric bacteria. Following reports that a mutated form of the NtrC protein may be responsible for the constitutive expression of the *dctA* gene, we investigated the possibility that a wild-type NtrC protein may also activate the *dctA* promoter. This work was carried

out in the heterologous host *E. coli* in which the *R. meliloti dctBD/dctA* regulatory pathway has been reconstituted. The *E. coli* NtrBC system was found to have both positive and negative regulatory effects on *dctA* expression (22). Under nitrogen-limiting conditions, NtrC could activate transcription of *dctA* even in the presence of the uninduced DctBD proteins. Under high nitrogen, NtrB instead repressed DctD-dependent transcription of *dctA* possibly by dephosphorylating DctD. However comparable cross-talks effects between the homologous NtrBC and Dct systems in *R. meliloti* could not be observed. This may indicate that the NtrBC system in *R. meliloti* has diverged in evolutionary terms such as cross-talk with the DctBD system (which does not exist in *E. coli*) cannot occur.

2.2. The effect of carbon source on dctA expression:

The presence of galactose as an additional carbon source was found to specifically repress the induction of *dctA* by dcAs in *R. meliloti* (27). The precise mechanism involved in this effect has yet to be elucidated.

We demonstrated that, in *E. coli*, the cAMP receptor protein, CRP, which normally acts on σ70-dependent promoters, has a repressive effect on *dctA*, a σ54-dependent promoter (17). We have found that CRP most likely prevents DctD from occupying its binding sites at the *dctA* promoter (30). The fact that CRP can interact with more than one form of RNA polymerase (Eσ54 and Eσ70) is of great interest as σ54 shares no significant homology with the σ70 family. We are now investigating whether *R. meliloti* has a CRP-like molecule and whether the activity of this molecule is regulated by cAMP levels.

3. Nitrogen metabolism and symbiotic nitrogen fixation:

Whereas free-living rhizobia assimilate combined nitrogen to support their own growth, the symbiotic forms do export fixed nitrogen to the host plant by a mechanism that remains to be determined. In addition, because nitrogen fixation is a very energy expensive process, it was anticipated that the plant nitrogen status may somehow influence bacteroid nitrogen fixation gene expression or activity. Accordingly, when increasing concentrations of combined nitrogen are added to a plant inoculated with *Rhizobium*, the symbiosis becomes less and less efficient and even, at high nitrogen concentrations, no nodules are observed.

3.1 Bacterial nitrogen metabolism (Cork, Dublin, Naples):

At the beginning of the project very little was known on the regulation of nitrogen metabolism in *Rhizobium*. We characterized components of the nitrogen assimilation pathway in two different biovars of *R. leguminosarum: R. l. viciae* and *R.l. phaseoli*. Two glutamine synthetases, GSI and GSII, have been demonstrated in rhizobia. GSI is similar to the single GS of enteric bacteria: it is a polymeric enzyme, relatively heat stable, made of 12 identical subunits (50 kDa each) and can be adenylated. In contrast, GSII is made up of 8 subunits of 36 kDa, which are heat labile (6). A *Rhizobium* gene, *glnT*, coding for a third glutamine synthetase, GSIII, has been identified and sequenced (4).

The *glnB* gene, coding for the PII regulatory protein, is located upstream of *glnA* in *R. leguminosarum* as in *B. japonicum*. A σ54-recognized promoter adjacent to *glnB* requires NtrC for activation of *glnBglnA* expression while a second promoter adjacent to glnA is neither σ54- nor NtrC-dependent. The concentration of the *glnBglnA* transcript in *R.l. viciae* was 2- to 3-fold higher in nitrate-grown compared to ammonium-grown cultures. The *glnA*-only transcript is not regulated by fixed N sources. The *ntrC* gene is also required for full expression of *glnII* (9): the *glnII*

transcript (coding for GSII) is enriched 10-fold in nitrate-grown as compared to ammonium-grown cultures.

We have identified the *ntrBC* genes of *R. leguminosarum* downstream of and cotranscribed with an ORF showing homology to an ORF located upstream of *ntrBC* in *R. capsulatus* and to the ORF1 located upstream of the *fis* gene of *E. coli*. The *R.l. phaseoli* ORF-*ntrBC* operon is negatively regulated by NtrC but does not require NtrC for expression (14). It is thus possible that, unlike in enteric bacteria, only changes in the activity state of NtrC, and not in its intracellular concentration, are important for activation of *glnB* and *glnII*.

We have shown that the absence of PII in *R.l. viciae* results in fully active (and presumably fully phosphorylated) NtrC (18). We have also shown that the PII protein of *R.l. viciae* is uridylylated in response to change from high N (ammonium) to low N (nitrate) conditions (12). Uridylylation was detected by labelling the cultures grown under various conditions with α^{32}P-UTP. The labelling was achieved by permeabilization of the cells with CETAB; crude extracts were immunoabsorbed and then run on SDS-PAGE. It is inferred that, as in *K. pneumoniae* and other enteric bacteria, PII in its deurydylylated form is required for the NtrC-P phosphatase activity of NtrB.

We have described a repressive effect of NH_3 on *nifA* expression in free-living *R. meliloti* (8). Our working hypothesis is that this effect might be mediated by the NtrBC system. In conditions of nitrogen excess, NtrB would be in a dephosphorylating mode and, possibly, counteract phosphorylation of FixJ by FixL. On *dctA* expression instead, no NH_3 effect could be observed.

In symbiosis, fixed nitrogen is exported from the nitrogen-fixing bacteroids into the plant cytosol. The export is believed to occur in the form of ammonia, which is then assimilated by plant enzymes. We focussed on the mechanism by which ammonium is taken up by free living cells and the conditions that influence this active transport. The underlying hypothesis was that some of the mechanisms involved in ammonium uptake could also play a role in ammonium export during symbiosis or, alternatively, that ammonium import could be inhibited during symbiosis. We have identified an active and inducible ammonium uptake system in free-living *R. meliloti* that has three striking characteristics:

(i) it operates at low (0.5 mM) but not high (20 mM) concentrations of ammonium chloride
(ii) it requires the nitrogen regulatory genes *ntrA* and *ntrC* genes
(iii) unexpectedly, this transport system is not necessary for growth at low ammonium concentration as evidenced from the phenotype of *ntrA* and *ntrC* mutants (25).

Given that the active ammonium transport detected in this work is not induced at high ammonium concentrations, it is unlikely to operate in the nodule. More probably, ammonium transport in the nodule occurs by passive diffusion. The active transport system may rather play a role in ion exchange across the membrane, perhaps as a constituent of an antiport system. Identification of the genes for ammonium uptake is presently ongoing. The symbiotic phenotype of mutants will be evaluated.

3.2 Plant nitrogen metabolism (Toulouse):

Two approaches have been taken in an effort to understand how plant nitrogen metabolism interferes with symbiotic nitrogen fixation. Because the first step in the assimilation of fixed nitrogen by the plant partner is catalysed by plant glutamine synthetase, identification of *Medicago truncatula* GS genes and study of their

expression has been undertaken. Three members of the GS family have been identified, the expression of two of them is significantly enhanced in the nodule simultaneously with the appearance of leghemoglobin mRNA.

Sense or antisense strategies have been used with the goal to either enhance or decrease GS expression in the nodule and to analyse the consequences on nitrogen fixation. *M. truncatula* transgenic plants have been obtained that expressed the sense or antisense mRNAs. However these constructs proved inefficient in affecting the endogenous levels of GS. Other constructs are presently being made.

In conclusion it can be stated that significant progress have been achieved in the identification of the regulatory pathways that control nitrogen metabolism, specially in the microsymbiont. This should allow us to elucidate the genetic links between nitrogen metabolism and nitrogen fixation.

MAJOR SCIENTIFIC BREAKTHROUGHS

(1) *In vitro* oxygen-regulated expression of *R. meliloti nifA* and *fixK* gene expression. First example of a complete signal transduction pathway of this sort reconstituted *in vitro*.
(2) Evidence for a functional alternative phosphorylation mechanism in the *R. meliloti* FixLJ two-component system.
(3) Identification of oxygen regulatory proteins in *R. leguminosarum* homologous to those in *R. meliloti*. Indications that the regulatory pathways are different in both species.
(4) Identification of fixN in Agrobacterium: its involvement in a microaerobically-induced oxidase activity.
(5) A number of novel, unexpected, regulatory effects on the *dctA* promoter, in both the homologous (galactose effect) and heterologous (Crp and Ntr effects) hosts.
(6) The CRP effect was the first observation of this σ70 regulator affecting a σ54-dependent promoter.
(7) The *E. coli* model for uridylylation of PII applies to a *Rhizobium* species. The amount and uridylylation state of PII, which depends on the glutamine/α-ketoglutarate ratio, reflects the amount of *glnB* and *glnII* transcripts present in high N-or low N-grown cultures.
(8) The permeabilized cell labelling technique is probably applicable to the analysis of other bacterial enzymes thought to modify proteins, without purification of the proteins involved.
(9) A novel function, ammonium uptake, has been identified as being under the control of nitrogen regulatory genes in *R. meliloti*.

MAJOR COOPERATIVE LINKS

4 meetings of the contractants (Paris 1991, Naples 1991, Toulouse 1992, Dourdan 1994). Visit of M. Iaccarino to Toulouse, visit of P. Boistard to Aachen. Exchange of 2 scientists between Bielefeld and Cork. Permanent exchange of materials and informations between the contractants.

PUBLICATIONS

Joint publications

1. *Colonna-Romano S., Arnold W., Schlüter A., Boistard P., Pühler A. and Priefer U.B. (1990) An Fnr-like protein encoded in *Rhizobium leguminosarum* biovar *viciae* shows structural and functional homology to *Rhizobium meliloti* FixK. Mol. Gen. Genet. **223**: 138-147.
2. *de Philip P., Boistard P., Schlüter A., Patschkowski T., Pühler A., Priefer U., O'Gara F., Boesten B. and Noonan B. (1992) Developmental and metabolic regulation of nitrogen fixation gene expression in *Rhizobium meliloti*. Can. J. Microbiol. **38**: 467-474.

Individual publications

3. Baev N., Amar M., Defez R. and Iaccarino M. (1992) The expression of the *nodD* and *nodABC* genes of *Rhizobium leguminosarum* is not regulated in response to combined nitrogen. FEMS Microbiol. Letts. **97**: 205-208.
4. Chiurazzi M., Meza R., Lara M., Lahm A., Defez R., Iaccarino M. and Espin G. (1992) The *Rhizobium leguminosarum* biovar *phaseoli glnT* gene. Gene **119**: 1-8.
5. de Philip P., Soupène E., Batut J. and Boistard P. (1992) Modular structure of the FixL protein of *Rhizobium meliloti*. Mol. Gen. Genet. **235**: 49-54.
6. Manco G., Rossi M., Defez R., Lamberti A., Percuoco G. and Iaccarino M. (1992) Dissociation of enzymatic activities of glutamine synthetase II after NH_4Cl treatment. J. Gen. Microbiol. **138**: 1453-1460.
7. Moreno S., Patriarca E., Chiurazzi M., Meza R., Defez R., Lamberti A., Riccio A., Iaccarino M. and Espin G. (1992) Phenotype of a *Rhizobium leguminosarum ntrC* mutant. Research in Microbiology **143**: 161-171
8. Noonan B., Motherway M. and O'Gara F. (1992) Ammonia regulation of the *Rhizobium meliloti* nitrogenase structural and regulatory genes under free-living conditions: Involvement of the *fixL* gene product? Mol. Gen. Genet. **234**: 423-428.
9. Patriarca E.J., Chiurazzi M., Manco G., Riccio A., Lamberti A., De Paolis A., Rossi M, Defez R. and Iaccarino M. (1992) Activation of the *Rhizobium leguminosarum glnII* gene by NtrC is dependent on upstream DNA sequences. Mol. Gen. Genet. **234**: 337-345.
10. Schlüter A., Patschkowski T., Unden G. and Priefer U. (1992) The *Rhizobium leguminosarum* FnrN protein is functionally similar to *Escherichia coli* Fnr and promotes heterologous oxygen-dependent activation of transcription. Mol. Microbiol. **622**: 3395-3404.
11. Waelkens F., Foglia A., Morel J.B., Fourment J., Batut J. and Boistard P. (1992) Molecular genetic analysis of the *Rhizobium meliloti fixK* promoter: identification of sequences involved in positive and negative regulation. Mol. Microbiol. **6**: 1447-1456.
12. Colonna Romano S., Patriarca E.J., Amar M., Bernard P., Manco G., Lamberti A., Iaccarino M. and Defez R. (1993) Uridylylation of the PII protein in *Rhizobium leguminosarum*. FEBS Letters **330**: 95-98.
13. Fitzmaurice A.M. and O'Gara F. (1993) A *Rhizobium meliloti* mutant, lacking functional γ-aminobutyrate (GABA) bypass, is defective in glutamate catabolism and symbiotic nitrogen fixation. FEMS **109**: 195-202.
14. Patriarca E.J., Riccio A., Taté R,. Colonna-Romano S., Iaccarino M. and Defez R. (1993) The *ntrBC* genes of *Rhizobium leguminosarum* are part of a complex operon subject to negative autoregulation. Mol. Microbiol. **9**: 569-577.
15. Reyrat J.M., David M., Blonski C., Boistard P. and Batut J. (1993) Oxygen-regulated *in vitro* transcription of *Rhizobium meliloti nifA* and *fixK* genes. J. Bacteriol. **175**: 6867-6872.
16. Schlüter A., Patschkowski T., Weidner S., Unden G., Hynes M.F. and Priefer U.B. (1993) Functional and regulatory characteristics of FnrN, an oxygen-responsive transcriptional activator in *Rhizobium leguminosarum* bv. *viciae*. In: Palacios 13. R., Mora J. and Newton W.E. (eds), New Horizons in Nitrogen Fixation. Kluwer Academic Publishers, Dordrecht, the Netherlands, p. 493.
17. Wang Y.P., Giblin L., Boesten B. and O'Gara F. (1993) The *Escherichia coli* cAMP receptor protein (CRP) represses the *Rhizobium* meliloti *dctA* promoter in a cAMP-dependent fashion. Mol. Microbiol. **8**: 253-259.
18. Amar M., Patriarca E.J., Manco G., Bernard P., Riccio A., Lamberti A., Defez R. and Iaccarino M. (1994) Regulation of nitrogen metabolism is altered in a *glnB* mutant strain of *Rhizobium leguminosarum*. Mol. Microbiol. **11**: 685-693.
19. Kennedy C., Doetsch N., Meletzus D., Patriarca E.J., Amar M. and Iaccarino M. (1994) Ammonium sensing in nitrogen fixing bacteria: functions of the *glnB* and *glnD* gene products. Plant and Soil **161**: 43-57.
20. Reyrat J.M., David M., Batut J. and Boistard P. (1994) FixL of *Rhizobium meliloti* enhances the transcriptional activity of a mutant FixJD54N protein by phosphorylation of an alternate residue. J. Bacteriol. **176**: 1969-1976.
21. Taté R., Patriarca E.J., Riccio A., Defez R. and Iaccarino M. (1994) Development of *Phaseolus vulgaris* root nodules. Molecular Plant-Microbe Interactions, in press.

Manuscripts submitted/in preparation:

22. Allaway D., Boesten B. and O'Gara F. NtrBC-dependent expression from the *Rhizobium meliloti dctA* promoter in Escherichia coli. (submitted).

23. Baev N., Abraham M., Riccio A., Iaccarino M. and Defez R. Inactivation of the *Rhizobium leguminosarum glnA* gene leads to a competition defective phenotype. (in preparation).
24. Batut J. and Boistard P. Oxygen control in *Rhizobium*. (submitted).
25. Devine M. and O'Connell M. Active ammonium uptake in *Rhizobium meliloti* is *ntrA* and *ntrC* dependent. (in preparation).
26. Galinier A., Garnerone A.M., Reyrat J.M., Kahn D., Batut J. and Boistard P. Phosphorylation of the *Rhizobium meliloti* FixJ protein induces its binding to a compound regulatory region at the *fixK* promoter. (submitted).
27. Motherway M., Archdeacon J., Pentony M., Boesten B. and O'Gara F. A repressive effect of galactose on *Rhizobium meliloti* C4-dicarboxylate transport gene (*dctA*) expression. (in preparation).
28. Patriarca E.J., Colonna-Romano S., Riccio A., Defez R. and Iaccarino M. DNA binding activity and phosphorylation state of the transcriptional regulator NtrC are influenced by the nitrogen source. (in preparation).
29. Schlüter A., Rüberg S., Weidner S., Krämer M. and Priefer U.B. (1994) A homologue of the *Rhizobium meliloti* nitrogen fixation gene *fixN* is involved in the production of a microaerobically induced oxidase activity in *Agrobacterium tumefaciens*. (submitted).
30. Wang Y.P., Kolb A., Buck M., Boesten B., O'Gara F. and Buc H. CRP interacts from long-distances with promoter-bound sigma 54 polymerase and blocks transcriptional activation of the *dctA* promoter. (submitted).
eX36B

Tomato transposon tagging: isolation of genes involved in disease resistance, hormone biosynthesis and plant cell development (BIOT CT-900192)

COORDINATOR:
J. HILLE, VU Amsterdam, Amsterdam, NL
PARTICIPANTS:
K. THERES / P. STARLINGER, Universität zu Köln, Köln, D
J. JONES, Sainsbury lab, Norwich, GB
I. TAYLOR, University of Nottingham, Nottingham, GB
R. DANKERT, RZ Research, Metslawier, NL

BACKGROUND INFORMATION

The challenge of this project can be simply stated. How can we isolate from plants, specifically tomato, the genes which correspond to interesting genetic traits, which have subtle or powerful effects in the whole plant, but which encode unknown proteins. Two main approaches are being developed for this purpose: cloning on basis of map position and cloning on basis of mutant phenotype (e.g. transposon tagging). In this coordinated program we pooled our resources in order to develop transposon tagging strategies in tomato and to apply these for cloning genes involved in disease resistance, hormone biosynthesis and plant cell development.

OBJECTIVES AND PRIMARY APPROACHES

The joint project is to develop a tomato transposon tagging system in order to enable each group to clone different genes. To this end it is aimed to jointly obtain a series of plants each carrying a transposable element insertion (either Ac or a derivative) in different positions that are equally spaced over the twelve chromosomes of tomato and to demonstrate the feasibility of this series for cloning specific genes by targeted transposon tagging. Within this project it is aimed to

(1) introduce Ac/Ds containing T-DNA's in tomato,
(2) to map T-DNA and transposon positions in a RFLP analysis,
(3) to analyse Ac/Ds transposition behaviour and
(4) to carry out targeted and non targeted transposon tagging experiments.

RESULTS AND DISCUSSION

All groups have introduced transposon containing T-DNA vectors into tomato by *Agrobacterium* mediated gene transfer. Both the autonomous Ac element and non-autonomous Ac-derivatives (Ds) have been positioned in various antibiotic resistance genes to permit selection for excision. Reintegration of Ds after excision is indicated by different marker genes inserted into the element itself. In these experiments also a novel cell autonomous selectable marker based on fusing a chloroplast targeting sequence to a spectinomycin resistance gene (aadA) has been developed. Alltogether, on all the twelve different chromosomes of tomato transposable element inserts have been mapped.

Typical germinal excision frequencies are usually around 10%. In the presence of Ac both excision of Ds from the original position and integration on new chromosomal locations could be established. For seven T-DNA loci Ds transposition is mainly to linked sites, however, for three T-DNA loci Ds transposition was

mainly inter-chromosomally. Because of the observed locus to locus variation in the transposition pattern it might be advantageous to study the pattern of transposition from the respective donor site prior to initiate large scale targeted transposon tagging experiments.

In non-targeted transposon tagging experiments several mutant phenotypes have been obtained that are presently being analysed. For a number of mutant phenotypes it was shown that there was no close linkage between the transposon insertion and the mutation. This indicates that there is a background in obtaining transposon induced mutations. However, at least a mutant in embryo development was shown to be tagged by the transposon (Jones lab).

Targeted transposon tagging experiments were aimed at the genes *Lanceolate* and *lateral suppressor* (Köln), *aurea* (Nottingham), the fungal disease resistance locus Asc (Amsterdam), a *dwarf* gene and the fungal disease resistance locus *Cf-9* (Norwich).

Two Ds containing T-DNA insertions located about 20 cM from the *Lanceolate* locus and about 40 cM from *lateral suppressor* were used for these experiments. F1 plants harbouring the T-DNA insertions were crossed with both mutants. Among 20,000 (*Lanceolate*) and 55,000 (*lateral suppressor*) progeny plants screened the respective mutant phenotypes were not observed. Both experiments suffered from the fact that no lines carrying a transposable element insertion close to one of these two genes had been obtained.

A line harbouring an Ac element 16 cM proximal to *aurea* was used in a tagging experiment to isolate aurea. Six putative Ac induced *aurea* mutants were obtained. In one of these an Ac copy was shown to co-map with *aurea*.

Using a line harbouring an Ac element located 15 cM from the *Asc* locus (donated by the Jones lab) a targeted tagging experiment for *Asc* was carried out. After screening 20,000 seedlings 3 mutants susceptible to infection by the fungus *Alternaria alternata* were isolated. For two mutants it could be shown that the mutated *Asc* locus did not co-segregate with the Ac element. The third mutant is more complicated, since in this plant Ac has amplified to 6 copies.

An Ac containing T-DNA was mapped 1 cM from the locus *dwarf*. From a targeted tagging experiment two *dwarf* mutants were obtained. The Ac insertions in these independent mutants map within 1 kb. of each other. Using a line containing a Ds element 3 cM away from the *Cf*-9 locus (donated by the Hille lab) targeted tagging experiments were carried out. A large number of mutants was obtained and it could be shown that 14 Ds insertions map in a 3 kb region. Currently, DNA sequence analysis of this region is in progress.

MAJOR SCIENTIFIC BREAKTHROUGHS

The set up of this coordinated program to establish transposon tagging technology in tomato was a major success. The goal to show the feasibility of isolating tomato genes by transposon tagging has fully been reached. Without BRIDGE support and such a coordinated program a major breakthrough, as reported here, would almost certainly not have happened.

MAJOR COOPERATIVE LINKS

Five meetings have been organised in this program, twice in Amsterdam and once in Norwich, Köln and Nottingham, to exchange the results of the participants. All groups have exchanged on a very regular basis both information, plasmid constructs and mapped tomato genotypes.

PUBLICATIONS

Biezen, E.A. van der, Haaren, M.J.J. van, Overduin, B., Nijkamp, H.J.J., and Hille, J. (1994). Heterologous transposon tagging as a tool for the isolation of plant genes. In Plant Molecular Biology Manual. S.B. Gelvin and R.A. Schilperoort, eds. Kluwer Academic Publishers, K2: 1-16.

Biezen, E.A. van der, Overduin, B., Nijkamp, H.J.J., and Hille, J. (1994). Integrated genetic map of tomato chromosome 3. Tomato Genetics Cooperative (in press)

Carroll, B.J., Klimyuk, V.I., Thomas, C.M., Bishop, G.J., and Jones, J.D.G. (1994). Germinal transpositions of the maize element Dissociation from T-DNA loci in tomato. Genetics (in press)

Haaren, M.J.J. van, and Ow, D.W. (1993). Prospects of applying a combination of DNA transposition and site-specific recombination in plants: a strategy for gene identification and cloning. Plant Mol. Biol. 23: 525-533

Hille, J., Rudenko, G., Biezen, E. van der, Blok, K., Overduin, B., Rommens, C., Nijkamp, J., and Haaren, M. (1993). Transposon induced chromosomal rearrangements. In Molecular biology of tomato. J.I. Yoder, ed. (Lancaster, PA: Technomic Publishing), pp. 93-106.

Jones, J.D.G., Jones, D.A., Balint-Kurti, P., Scofield, S., Dickinson, M.J., Thomas, C.M., Harrison, K., Bishop, G., Ashfield, T., Hammond-Kosack, K., and Carroll, B. (1993). Towards an understanding of the function of tomato genes for resistance to *Cladopsorium fulvum*. In Molecular biology of tomato. J.I. Yoder, ed. (Lancaster, PA: Technomic Publishing), pp. 263-274.

Knapp, S., Larondelle, Y., Robberg, M., Furtek, D., and Theres, K. (1994). Transgenic tomato lines containing Ds elements at defined genomic positions as tools for targeted transposon tagging. Molecular & General Genetics **243**: 666-673.

Scofield, S.R., Jones, D.A., Harrison, K., and Jones, J.D.G. (1994). Chloroplast targeting of spectinomycin adenyltransferase provides a cell autonomous marker for monitoring transposon excision in tomato and tobacco. Molecular & General Genetics **244**: 189-196.

Thomas, C.M., Jones, D.A., English, J.J., Carroll, B.J., Bennetzen, J.L., Harrison, K.A., Burbidge, A., and Jones, J.D.G. (1994). Analysis of the chromosomal distribution of transposon-carrying T-DNAs in tomato using the inverse polymerase chain reaction. Molecular and General Genetics **242**: 573-585.

Identification of regulatory genes controlling major metabolic pathways (BIOT CT-900164)

COORDINATOR:
Michel CABOCHE, Lab. Biologie Cellulaire, INRA, Versailles, F
PARTICIPANTS:
1. C. TONELLI, Univ. studi di Milano, Milano, I
2. G. GIULIANO, ENEA, Roma, I
3. C. MARTIN, John Innes Institute, Norwich, GB
4. W.D.O. HAMILTON, A.G.C., Cambridge, GB
5. J. PAZ-ARES, CSIC, Madrid, E
6. R.M. WALLSGROVE, Rothamsted Experimental Station, GB
7. R. DELON, SEITA, Bergerac, F

BACKGROUND INFORMATION

Metabolic pathways are under the control of a number of regulatory elements. Such genes should be identified and deregulated in the view of crop improvement.

OBJECTIVES AND PRIMARY APPROACHES

The regulatory genes from the anthocyanin biosynthesis pathway and the nitrate assimilatory pathway were identified by genetic or molecular approaches, cloned and characterized at the molecular level by gel retardation assays, and transcription studies. The deregulated expression of the corresponding pathways by genetic engineering and the characterization of the obtained transgenic plants has been the goal of this project.

RESULTS AND DISCUSSION

I. Versailles

A GUS reporter sequence was transcriptionally linked to the promoter of the *nia-1* gene of *N. tabacum*. This fusion was functional in transient expression assays done with protoplasts derived from mesophyll cells of *N. tabacum*. However none of the regulations known to affect steady state levels of the *nia-1* transcript were detectable under these experimental conditions. Particule gun bombardment was then used to study the transient expression of a *nia1*-GUS reporter gene in mesophyll tissues of plantlets with similar conclusions. Transgenic plants carrying either this fusion or translational fusions of GUS linked to the promoter of either the *nia-1* or *nia-2* gene of *N. tabacum* were obtained by *Agrobacterium*-mediated transfer. A low proportion of them expressed appropriately GUS. The accumulation of the mRNA significatively increased in a NR-deficient background indicating a regulation by N-metabolites and by nitrate.

PCR approaches and cross-hybridization techniques were used to isolate respectively a partial cDNA from *N. plumbaginifolia* and a full length cDNA from *N. tabacum*. These clones encode NTL1, a NIT2-like protein, characterized by a single zinc finger domain, defined by the $C\text{-}X_2\text{-}C\text{-}X_{18}\text{-}C\text{-}X_2\text{-}C$ amino acids and associated with a basic region. The amino acid sequence of NTL1 is 60% homologous to the NIT2 sequence in the zinc finger domain.The characteristics of the *Ntl1* gene expression are compatible with that of a regulator of the nitrate assimilation pathway, namely a faint nitrate inducibility and a regulation by light.

The NTL1 protein has features typical of a transcription factor, including nuclear targeting signals and activator domains. Wild type *N. plumbaginifolia* plants have been transformed with constructs containing either the entire or the zinc finger and basic region of *Ntl1* cDNA, driven in the antisense orientation by the 35S promoter. No dramatic effect was observed except for a high level of sterility and an overexpression of a glutamine synthetase gene in three independant transformants. However, analysis of their progeny has failed to demonstrate any correlation between this overexpression and the presence of the constructs.

II. Milano

a) Expression of regulatory and structural genes of anthocyanin pathways in maize.

1) Tissue specific expression: we have investigated the tissue and cell specific localization of different alleles of the R family. Sn or R homologous transcripts are specifically expressed only in the tissues that are under the gene control. Presence of R or Sn transcripts correlates precisely with the induction of A1 and C2 messengers and with the accumulation of the anthocyanin pigment. Absence of a gene of the R family, as in the r-g genotype, shows, in all the tissues examined, a lack of R, A1 and C2 transcripts correlating with colorless phenotypes. These results confirm that the pattern of pigmentation is strictly dependent on the allelic combination of the genes of the R family.

2) Effect of different light qualities: Pericarp and aleurone anthocyanin accumulation is controlled respectively by Sn and R, allowing a dramatic increase in pigmentation in response to white light irradiation. Irradiation of these two tissues at succeeding developmental times during seed formation and in different genotypes has disclosed that the two tissues differ in their response to light during their development;

(i) irradiation of developing seeds with light of different quality, but equal fluence rate, indicates that both red and blue light elicit a positive response in the aleurone (under R control) while in the pericarp (under Sn control) only blue light is effective;

(ii) presence of R and Sn together in the genome leads to a drastic reduction in pericarp pigmentation in comparison to that observed in r Sn genotypes. White and blue light treatments lead to an enhanced transient accumulation of Sn mRNA in the pericarps with a maximum level of induction after 18-24 h of illumination. C2 and A1 transcripts increased in a coordinated manner after 24 h of illumination as expected, considering the regulatory role played by Sn. R transcript was not affected by light treatment suggesting that R is controlled at the post-transcriptional level.

b) Sequence analysis and comparison of cDNA and promoter clones

We have completed the whole cDNA sequence of the Sn:bol3 allele and the comparison with the cDNA sequences of Lc and of the S component of the R complex, referred as R-S. It contains a 616 aminoacids longORF. In the leader sequence two short open reading frames of 38 and 15 amino acids are also present. The putative protein contains a large acidic domain and a basic region part of which was found to have similarity to the myc family of oncogene proteins. This region has been proposed as a novel DNA binding and dimerization motif called helix-loop-helix (HLH). A comparison between the three cDNA sequences of Sn, Lc and R-S discloses a very high degree of homology particularly in the translated region. These findings support the hypothesis that Sn, Lc and R genes encode functionally

related proteins. The distinct pattern of pigmentation determined by these genes may thus reflect differences in the promoter region rather than functional differences in their product.

We have obtained genomic clones containing promoter sequences of three alleles, all showing different tissue specific expression. The comparison of these sequences highlights that r-r promoter is identical to that of Sn, while that of rch-hopi shows several polymorphisms.

c) Interaction among members of the R family

R genes extracted from R/Sn heterozygotes, exhibit a Sn phenotype in their mesocotyls, as if in the heterozygous association Sn had imposed an imprinting on the R gene leading to its expression in an additional tissue. This imprinting is negatively affected by chromosomal rearrangements and varies in its frequency (10-40%) according to the R alleles. Molecular analysis (RACE) points to the activation at the transcriptional level of the imprinted gene in a tissue where it is normally repressed. Recent studies on transgenic plants have disclosed similar events taking place between ectopic genes and the resident homologous genes in ways that lead to changes in gene expression at either one or both locations.

Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.)

Clones coding for phenylalanine ammonia-lyase (PAL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydoxylase (F3H), dihydroflavonol 4-reductase (DFR), leucoanthocyanidin dioxygenase (LDOX) and UDP Glucose: flavonoid 3-O -glucosyl transferase (UFGT), were isolated by screening a cDNA library, obtained from mRNA from seedlings grown in light for 48 hours using snapdragon (*Antirrhinum majus*) and maize heterologous probes. A cDNA clone coding for stilbene synthase (StSy) was isolated by probing the library with a specific oligonucleotide. The percentages of sequence similarity ranged from 65% (UFGT) to 90% (CHS and PAL). The analysis of the genomic organization and expression of these genes in response to light shows that PAL and StSy genes belong to large multigene families, while the others are present in one to four copies per haploid genome. The steady state level of mRNAs encoded by the flavonoid biosynthetic genes as determined in young seedlings is coordinately induced by light, except for PAL and StSy, that appear to be constitutively expressed.

III. Roma

a) Cloning and characterization of bZIP transcription factors binding the G box of tomato light-regulated genes.

Using PCR with degenerated oligonucleotides, several cDNAs homologous to G box binding factors from other plants have been cloned. One of them encodes a 45 kDa basic leucine zipper (bZIP) factor, termed GBF1. When expressed in E. coli, this factor binds specifically the G box, although with a sequence specificity different from that of the GBF activity from leaf extracts. GBF1 differs from tomato leaf GBF also in the fact that its binding to DNA is not controlle by phosphorylation with casein kinase II. In transient expression assays, GBF1 is not able to trans-activate a full-length form af the rbcS-3A promoter, while it represses 4-6 fold a truncated, but still active, form of the same promoter (Ponti et al., unpublished). We are working on two alternative hypotheses:

i) that GBF1 needs cooperation with proteins binding on nearby sequences to exert its trans-activation.
ii) That it is itself a transcriptional repressor. The high level expression of GBF1 mRNA in fruits is in accordance with the latter hypothesis.

b) Characterization of transcription factors binding the I box of tomato light-regulated genes.

We have completed the characterization of the activities from tomato crude nuclear extracts which bind the I box, a regulatory motif found in the plant light-regulated rbcS, cab and nia promoters which contains a GATAA(A/G) core sequence. Due to this fact, the I box is a possible target for the binding of members of the GATA family of zinc finger transcription factors. We have received from the Versailles group and expressed in *E. coli* the cDNA encoding NTL-1, a tobacco GATA family factor. The NTL-1 protein shows high affinity for the I box of the rbcS gene, but not the G box from the same gene of the I boxes from the cab or nia genes. Therefore, NTL-1 shows characteristics similar to one of the activities previously characterized in crude extracts, belonging to the IBF-1 group.

c) Study of the light-regulation of carotenoid biosynthesis genes.

The PDS promoter has been introduced in transgenic tobacco plants. It is induced by photooxidative stress, and in etiolated seedlings. We are trying to localize the cis-acting sequences and trans-acting factors responsible for this type of regulation.

IV. Norwich

Our genetic approach has been through the identification of mutations in genes regulating anthocyanin biosynthesis, an analysis of their regulatory effects on structural genes, and isolation of the genes involved where possible. We have made most progress with the *mixta* gene which enhances intensity of floral pigmentation in the corolla lobes. A somatically unstable allele of this gene was identified and cloned. The *mixta* gene encodes a Myb-related transcription factor, most similar to the *myb*-related to the *myb*Ph1 gene product identified by the Madrid group in *Petunia*. However, the *mixta* gene product does not appear to regulate transcription of anthocyanin biosynthetic genes to enhance pigment intensity. Despite the difference in the visual apparence of the mutant flowers, extracts show the same amount and types of anthocyanin to be present in *mixta* and wild- type flowers. The expression of the structural genes is also the same. The diffrence in pigment intensity appears to depend on the specialised shape of petal epidermal cells. These cells are conical in wild-type, but flat in *mixta* lines. The major effect of *mixta* appears to be on the production of specialised epidermal cell wall. We have therefore established that a transcription regulator controls specialised cell wall production, so linking the control of anthocyanin biosynthesis to control of different metabolic processes. At present we are collaborating with the group in Madrid to introduce *myb*Ph1 in antisense into *Petunia* to establish whether it has the same function as *mixta* We are also expressing *mixta* in *Petunia* to examine the effects of its overexpression on epidermal cell differentiation.

Analysis of another mutant, *Eluta*, has demonstrated it affects the regulation of structural gene expression. Mutants have much lower levels of F3H, DFR, Caridi and UFGT expression but normal levels of CHS and CHI expression. This establishes that *eluta* regulates the anthocyanin biosynthetic pathways in a similar manner to *delila*. A collaborative screen of ems/transposon mutagenesised populations with Zsusanna Schwarz-Sommer of the Max Planck Institut, Köln (BRIDGE Programme BIOT CT-90171), identified several new *Eluta* mutants, some of which were somatically unstable, indicating the phenotype of the mutant allele, and also

providing material for transposon tagging strategies. We have pursued our study of the mode of action of the *delila* gene by preparing antiserum for protein produced by expression of the cDNA in *E. coli*. We have also tested whether *delila* will activate structural gene expression in yeast using a reporter gene driven by the promoter fragments of the targets gene. Our results suggest that *delila* will not transactivate gene expression alone, implying that it is involved in co-oprative control of anthocyanin biosynthesis.

We have followed a similar strategy in our study of six cDNA clones to Myb-related transcription form *Antirrhinum* flowers. Antiserum prepared to one expressed specifically in flowers, *myb*305, has been used to identify a DNA binding factor in tobacco and *Antirrhinum* nuclear extracts that binds to PAL and CHS gene promoters. From these experiments we have established that *myb*305 can activate some structural genes of the general phenylpropanoid and anthocyanin biosynthetic Our study of the function of the other Myb-related transcription factors through over-expression in transgenic tobacco has provided striking phenotypic changes in plants over-expressing Myb308. Our most recent results suggest that Myb308 represses phenolic acid biosynthesis. Phenolic acids are formed via a branch of the general phenylpropanoid pathway and are therefore related biochemically to anthocyanins by sharing a common precursor pathway. These results show that structurally-related transcription factors may be allocated control of different branches of related metabolic pathways.

We have also collaborated with the group in Milan on the isolation of structural genes anthocyanin biosynthesis in grape (see above).

V. Cambridge

Since anthocyanin is best visualised in tomato hypocotyls 3 week old tomato hypocotyl as starting material. A cDNA library was made from the tomato hypocotyl mRNA and screened using a mixture of degenerate oligonucleotides derived from the conserved region of the 6 *Antirrhinum myb* genes. Twenty positive clones were identified and terminal sequencing of both ends revealed that all the clones were *myb*-related, 14 of them being independent. Moreover, six clones were shown to be full length (for the coding region at least) by comparison with the *Antirrhinum myb* clones.

The expression patterns of these tomato *myb* genes were studied by Northern blot analysis using poly A^+ RNA from 7 different types of tissues (hypocotyl, leaf, root, green and red fruits, immature and mature flowers). A number of clones hybridized to multiple transcripts, some of which were tissue-specific. Other clones hybridized to double or triple transcripts, whereas some hybridized to only one transcript. The sizes of bands detected ranged from 0.4 to 5 kp in length. The multiple transcripts may result from expression from different promoters, termination at different polyadenylation sites of the use of alternative splice sites.

Having studied the expression patterns, three clones showing different expression patterns were fully sequenced. One clones was similar to the *Antirrhinum myb* gene 308 in that the first 22 amino acid residues are identical in both clones and two conserved regions outside the *myb* repeats were also found. The genomic organisation of the three clones were studied by Southern blot analysis. Two of them are single copy genes and one has multiple copies.

An antisense aproach was adapted in an attempt to elucidate the functions of these tomato *myb* genes. Antisense constructs were made for the three fully sequenced clones in a plant expression vector which contained the CaMV 35S promoter, NOS

terminator and a GUS reporter gene. The variable regions of the clones (outside the conserved *myb* DNA binding domain) were used. *Agrobacterium*-mediated tomato transformations were carried out. A total of 3 600 explants were inoculated with the three antisense constructs, from which only one transgenic plant was obtained. Unfortunately, this transgenic plant showed no visual differences to the wild type. Two of the constructs described were also used to transform potato. Three transformants were obtained for one of the constructs which showed some phenotypic changes. RFLP mapping was another method used to try to elucidate the functions of these tomato *myb* genes. This work, which is not yet complete, is being carrried out in collaboration with Dr P. Linhout (CPRO-DLO, Wageningen). Autoradiographs were obtained for eight of the tomato *myb* genes and will be sent to Dr Linhout for linkage analysis.

VI. Madrid

Using a binding site selection protocol, previously, we found that MYB.PH3 binds to the sequence A a a C G/C G T T A (MBSI), a type of site very similar to that of animal c-MYB. Further analysis of the sequences of selected nucleotides allowed the definition of a different sequence to which MYB.PH3 also binds with high affinity: G T T A G T T A (MBSII). This second sequence resembles that of other plant MYB proteins, such as the *Antirrhinum* Am305 and the maize P proteins. Thus the binding mechanisms operating in two classes of MYB proteins may coexist in a single MYB prtein, such as MYB.PH3. Hydroxi-radical footprinting studies indicated that MYB repeats. To evaluate whether MYB.PH3 can activate transcription we used the yeast system. For this purpose, two MBSI or MBSII sequences were places in front of the minimal *CYC1* promoter: *LacZ* gene. The activity of these reporters was shown to be 15-fold enhanced by the presence of MYB.PH3, resulting from expression of its cDNA under the control of the *GAL10* promoter. Thus, the type of binding site doesn't predetermine the transcriptional effect (activation or repression) of MYB.PH3 on its target genes. The tissular location of MYB.PH3 was examined by *in situ* immunolocation experiments which showed that this protein is located in the epidermal cells of *Petunia* petals. This cell type is highly active in flavonoid biosynthesis, in agreement with a role of MYB.PH3 in the regulation of flavonoid biosynthetic genes. Moreover, a computer assisted search revealed the presence of MYB.PH3 binding sites several flavonoid biosynthetic genes, particularly *chalcone synthase* genes. Transgenic plants overexpressing 'sense' or 'antisense' *myb.Ph3* RNAs are being produced in collaboration with the Norwich group.

VII. Rothamsted

Protocols have been succesfully developed for the selection of barley mutants with de-regulated expression of nitrate reductase (EC 1.6.6.1). The selection has been carrried out with seed populations derived from chemically-mutagenised material (sodium azide). To date 25,000 seedlings from the M2 generation have been screened, and 105 lines with greater than wild-type NR activity identified in the first screen. Of these, 88 survived to produce M3 seed, and to date 52 M3 seedings have been screened.

One of these lines, designated RNR 93/1, shows inheritance of the constitutive NR expression. Average *in vivo* NR activity in the leaves of this line is 7.7 nmole N02 produced per leaf slice per hour, compared to a fully-induced wild type activity of 27.9 and non-induced (no nitrate) wild-type activity of less than 1.0. Selfed M4 seeds from this line have been obtained and reciprocal back-crosses with the

parent line have been made. Initial analysis of M3 seedlings suggests that the mutation has no effect on the expression of nitrite reductase, which would indicate that it has not affected a common regulatory gene. Extractable *in vitro* NR activity is detectable in non-induced mutants, so the selection is not an artefact of the *in vivo* assay used. A complete biochemical and genetic analysis will be made (of this and other lines) in collaboration with Dr John Wray (University of St Andrews, Scotland) in a BBSRC-funded 3-year programme starting in May 1994.

Several of the M2 selections had very high NR activity, greater than double the fully-induced wild-type activity. None of these lines survived to produce seed, and all grew slowly and weakly, under a variety of growth conditions. Attempts to maintain the lines in cell culture also failed, and it is possible that such high NR expression (in the absence of nitrate) is itseld deleterious to the plant. A modified selection strategy, aimed at overcoming this problem, is in progress and will form part of the on-going project with Dr Wray.

The mutants selected in this project are the first NR regulatory mutants ever isolated in any species of higher plant. It is hoped that detailed analysis of them will lead to a fuller understandingof the regulatory mechanisms controlling NR expression in plants, and the identification of possible regulatory genes.

VIII. Bergerac

Plants overexpressing nitrate reductase were produced, using as an expression vector, a construct made by the Versailles group invoved in this project. In this construct, the NR cDNA of *Nicotiana Sylvestris* was placed under the control of the strong promoter 35 S of the Cauliflower virus. Four hundred transformed buds were obtained. Tobacco plants possessing the highest NR activity were screened. Four lines were selected from the cultivar PB D6 and one from the cultivar BB16. Field experiments were carried our in 1992 and 1993.

These trials had three purposes:

— To observe the evolution of NR activity during plant growth.
— To assess the effects of the overexpression of NR activity on the chemical composition of the leaves and growth.
— To evaluate the capacity to assimilate nitrate.

Transgenic plants have a phenotype similar to the wild type, but have less leaves (2 or 3) and they are smaller (10 centimeters in average). The floral induction and the anthesis occur 6 days apart. NR activity tends to remain at a higher level than in the WT. Although a steady decrease of NR activity during plant growth is observed for both types.

The metabolism of the transgenic plants is more active. In fact, the nitrate content except for a line, decreases strongly. For example, in the line 30.51. 2"PB D6 nitrate content decrease by 11 to 75% and overall 36%. This decrease of nitrate content is more important in the midrib where nitrate is generally accumulated in the wild type. The extra-requirement of reducing potential due to nitrate reduction causes a strong decrease in reducing sugars concentration for all the lines and all the leaf positions.

Some plants as the lines 30.51.2 PB D6 and 34.2.5. BB 16 exhibit a NR deficient phenotype characterized by a yellowing of the leaves. In these lines the introduced genes interact with homologous resident genes by a mechanism called co-suppression and leading to the loss of expression of all nitrate reductase coding sequences.

Field trial experiments for nitrogen utilization efficiency:

In 1993, the line 34.19.4PB D6 was cultivated with four rates of nitrogen fertilisation (0, 75, 200 and 400 kg/ha/N). The capacity of this line to grow with a low rate of nitrogen is better than for the wild type and the yield increases appreciably.

MAJOR SCIENTIFIC BREAKTHROUGHS

Tissue specific expression of transcriptional activators in maize; isolation and sequence of transcription factor promoters in maize, isolation, sequence and characterization of cDNA clones of seven structural genes of anthocyanins pathway in grape.

A single MYB protein, MYB.PH3, can bind to and activate transcription from, two different types of sequences. The presence of these sequences in the promoter region of flavonoid biosynthetic genes and the location of MYB.PH3 in the epidermis of *Petunia* petals, in which flavonoids are highly actively synthesized, strongly suggest that this MYB protein also regulates flavonoid biosynthesis.

MAJOR COOPERATIVE LINKS

In collaboration with the laboratory of G.Giuliano (ENEA, Roma), we have analysed the interactions between NTL1 and different GATA boxes. Preliminary results show that our GATA protein recognizes specifically sequences derived from the rbcs 3a promoter but not from the tomato nia promoter. NTL1 may therefore be involved in the control of the expression of light-regulated genes.

The group in Milano in collaboration with the group at John Innes have isolated and sequenced cDNA clones of seven structural genes of anthocyanins pathway in grape (Vitis vinifera) using maize and Anthyrrinium clones as heterologous probes.

The Norwich group has provided vector and strains for the transcription activation experiments in yeast. In addition, the Norwich group collaborates in the experiments with transgenic *Petunia* plants.

PUBLICATIONS

Joint publication

Sparvoli F., Martin C., Scienza A., Gavazzi G. and Tonelli C. (1994). Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera L*). Plant Molecular Biology. **24**, 743-755.

Individual publications

Avila J, Nieto C, Cañas L, Benito M J, Paz-Ares J (1993). *Petunia* hybrida genes relatred to the maise regulatory *C1* gene and to animal *myb* protooncogenes. The Plant Journal **3**, 553-562.

Borello, U., Ceccarelli, E. and Giuliano, G. (1993). Constitutive, light-responsive, and circadian clock-responsive factors compete for the different I box elements in plant light-regulated promoters. Plant J. **4**, 611-619.

G. Consonni, F. Geuna, G. Gavazzi, C. Tonelli (1993). Molecular homology among members of the R gene family. Plant J. **3**: 335-346

G. Consonni, A. Viotti, S. Dellaporta, C. Tonelli (1992). cDNA Nucleotide sequence of Sn, a regulatory gene in maize. Nucleic Acids Res. **20**: 373.

Daniel-Vedele F., Caboche M. (1993). A tobacco cDNA clone encoding a GATA-1 zinc finger protein homologous to regulators of nitrogen metabolism in fungi. Molecular and General Genetics **240**: 365-373

S. Dolfini, G. Consonni, M. Mereghetti, C. Tonelli (1993). Antiparallel expression of the sense and antisense transcripts of maize a-tubulin genes. Mol.Gen.Gen. **241**: 161-169)

G. Gavazzi, G. Consonni, C. Tonelli. R gene family in maize: A model system to monitor epigenetic changes in response to internal and external signals. Current Topics in Molecular Genetics. (in press)
Giuliano, G., Bartley, G.E. and Scolnik, P.A. (1993). Regulation of carotenoid biosynthesis during tomato development. Plant Cell **5**, 379-387.
Godon C., Caboche M., Daniel-Vedele F. (1993). Transient plant gene expression: a simple and reproducible method based on flowing particle gun. Biochimie **75**: 591-595
Jackson D, Culianez-Macia F., Prescott A., Roberts K. and Martin C. (1991) Expression patterns of Myb-genes from *Anthirrhinum* flowers. Plant Cell **3**: 115-125.
Jackson D., Roberts K., Martin C. (1992). Temporal and spatial control of anthocyanin biosynthetic genes during flower development in *Antirrhinum* majus. Plant Journal, **2**: 425-434.
Lin Q., Merryweather A., Hamilton W.D.O. (1994). Cloning and characterization of *myb* genes from tomato (Lycopersicon esculentum). In preparation.
Martin C. and Gerats T. (1993) Control of pigment biosynthesis genes during petal development. Plant Cell **5**: 1253-1264.
Martin C., Prescott A., MacKay S., Bartlett J. and Vrijlandt E. (1991). The control of anthocyanin biosynthesis in flowers of *Antirrhinum* majus. Plant Journal **1**: 39-45.
Noda K., Glover B.J., Linstead P., and Martin C. (1994). Flower colour depends on specialised cell shape controlled by a Myb-related transcription factor. Nature, (in press).
Sablowski R.W., Moyano E., Culianez-Macia F.A., Schuch W., Martin C. and Bevan M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO journal **13**: 128-137.
Vaucheret H., Marion-Poll A., Meyer C., Faure J.D., Marin E., Caboche M. (1992). Interest in and limits to the utilization of reporter genes for the analysis of transcriptional regulation of nitrate reductase. Molecular and General Genetics **235**: 259-268.

The plasma membrane and the tonoplast of plant cells as targets to increase plant productivity (BIOT CT-900175)

COORDINATOR:
A: S. DELROT, Univ Poitiers, Poitiers, F
PARTICIPANTS:
B: L. WILLMITZER, Genbiologische Forsch., Berlin, D
C: H. CANUT, Univ. Toulouse, Toulouse, F
D: E. MARRÉ, Univ. Milano, Milano, I
E: B. FORDE, Rothamsted Experimental. Stn., Herts, GB
F: A. SMITH, Univ. Oxford, Oxford, GB
G: G. SILVESTRI, COSPAG, Fidenza, I

MAJOR SCIENTIFIC BREAKTHROUGHS

Plant productivity depends on the activity of various membrane transporters which allow the distribution of assimilates from photosynthesizing leaves towards storage organs (fruits, roots, tubers) harvested for human consumption. The programm gathered the expertise of physiology, biochemistry, and molecular biology groups in order to identify and characterize some of these transporters.

OBJECTIVES AND PRIMARY APPROACHES

The initial objectives of the programm were to characterize several transporters (sucrose, amino acids, ions) of the plasma membrane and a major transporter (malate) of the tonoplast, and to clone the corresponding genes. The primary techniques developed include the use of purified plasma membrane and tonoplast vesicles, patch clamp, reconstitution of transporters, mutagenesis, complementation of yeasts and of *Xenopus* oocytes. Cloning of the transporters would provide a necessary basis for the manipulation of sink/source interactions in the plant, and hence for the control of productivity.

RESULTS AND DISCUSSION

I. The sucrose transporters of the plant plasma membrane

1. Sucrose efflux from the mesophyll cell

Sugar transport in the leaf involves two carrier-mediated steps for sugars, i.e. efflux from the mesophyll cell, and uptake into the phloem cells. Studies with plasma membrane vesicles from sugar beet leaf have shown that the sucrose carrier involved in passive sucrose efflux is different, or at least in a state different, from that involved in proton-driven active influx (Laloi *et al.*, 1993).

2. Evidence for the incorporation of a NEM-sensitive H^+/sucrose cotransport into the plasma membrane of leaf cells during the sink/source transition

A young leaf is still heterotrophic for carbon and imports its assimilates from mature leaves via a symplastic pathway. When the leaf matures, enough reduced carbon becomes available for export, and structural and ultrastructural changes allows its transport out of the leaf. Experiments with plasma membrane vesicles have shown that a NEM-sensitive H^+/sucrose cotransport process is incorporated (or activated) into the membrane during the sink/source transition. The protein

pattern analyzed by mono- and bi-dimensional gel electrophoresis shows that this is accompanied by the appearance of a new set of proteins in the 40 kD region.

3. Biochemical characterization of the H^+/sucrose cotransporter

Previous biochemical and immunological work had allowed the identification of an intrinsic 42 kD polypeptide of the plasma membrane as a putative sucrose transporter. Several new polyclonal sera and ascitic fluids directed against the 42 kD region of the plasmalemma have been raised in mice. These sera selectively inhibited proton-driven uptake of sucrose into purified plasma membrane vesicles (Gallet *et al.*, 1992). This confirmed, on a functional basis, the previous hypothesis that a 42 kD polypeptide is involved in sucrose transport across the plasma membrane.

An oriented expression cDNA library in λ gt 11 was constructed from mature sugarbeet leaves, but attempts to isolate the carrier gene by screening cDNA expression libraries with this antiserum have failed.

The 42 kD putative sucrose transporter previously identified had been characterized after denaturating gel electrophoresis. Successful attempts were made to purify the sucrose carrier in a functional state, using solubilization of the plasma membrane proteins by non-denaturating detergents and high performance liquid chromatography (gel filtration + ion-exchange, or affinity chromatography on a sucrose-sepharose column + ion-exchange). Fractions exhibiting the highest transport activity after reconstitution in proteoliposomes were enriched in two close bands at 42 kD. A 55 kD band was also identified by affinity chromatography, which recognize sucrose.

4. Regulation of sucrose transport

To study the regulation of sucrose transport in leaf tissues at the biochemical level, we have developped the model of cutting and ageing. Ageing is a well known process which consists in isolating leaf pieces and floating them, after removal of the lower epidermis, on a large volume of osmoticum for several hours. 'Cutting' consists in excising the leaf and dipping the petiole for 12 hours into water before preparation of the leaf discs or of the plasma membrane.

After cutting, sucrose uptake was doubled compared to fresh tissues, while the uptake of 3-O-MeG and valine remained unaffected. After ageing, there was a general stimulation of uptake, which represented 400% of the uptake measured in fresh discs for sucrose, 3-O-MeG and valine. The data show that cutting is sufficient to induce dramatic and selective changes in the uptake properties of leaf tissues, and that the effects of cutting and ageing on the uptake of organic solutes are clearly different.

Parallel experiments were run with purified plasma membrane vesicles prepared from **fresh** and **cut** leaves. The data show that

(a) the changes induced by cutting on sucrose uptake by leaf discs are due to membrane phenomena, and not to the metabolism of sucrose

(b) the study of sucrose uptake with plasma membrane vesicles give a good account of the physiological situation

(c) the specific effects induced by cutting on the sucrose uptake system are not lost during the preparation of the plasma membrane vesicles.

5. *Isolation and characterization of sucrose transporter genes by functional complementation of yeast mutants*

To isolate different genes based on their functional expression, the Berlin group developed complementation techniques . In order to set up an artificial complementation system, yeast expression vectors, cDNA libraries under control of yeast promoters and efficient transformation protocols were established. A yeast (*Saccharomyces cerevisiae*) strain deficient in secreted invertase, but able to metabolize ingested sucrose due to expression of a sucrose cleaving activity has been used successfully as an artificial complementation system to isolate a sucrose transporter cDNA from spinach and potato. The expression in yeast has allowed to analyze the biochemical properties of the transporters directly. The two proteins from spinach and potato are very similar concerning Km, pH-dependence, and inhibition of transport by protonophores, thiol modifying agents and DEPC. The Km value of the carriers for sucrose (about 1 mM), and the specificity towards others sugars is in good agreement with the data described for the sucrose carrier in plant systems.

The transporter genes from spinach (SoSUT1) and potato (StSUT1) encode highly hydrophobic proteins consisting of two sets of six membrane spanning regions, separated by a large cytoplasmic loop. Comparison of the potato and spinach carriers reveals 68% identity on the protein level. The areas with the highest variability are the N- and C-terminal extensions and the large central loop. Although no sequence homologies were found to the prototype of sugar transporting proteins, that is the lactose permease from *E. coli*, the transporters seem to be related not only in being disaccharide transporters but also in their structure with twelve membrane spanning regions separated by a large central loop. Interestingly, a conserved motif (RXGRR) is located in the second loop of the sucrose transporters. There is a synonymous position for this motif as in lactose permease, and was found also in several other transporters.

Correlative expression of the carrier gene with the development of active transport activity of maturing leaves suggests a role for StSUTI in phloem loading. StSUTI RNA is also found in stems though at a lower level as compared to leaves. This is taken as an argument against a sole function of the transporter in retrieval along the translocation pathway. By RNA *in situ* hybridization, the expression of the carrier was localized in the phloem thus supporting the role of the carrier in phloem loading.

To analyze directly the function of the transporter, potato plants were transformed with an antisense gene under the control of the CaMV 35S promoter. Antisense inhibition of the sucrose transporter leads to drastic phenotypic effects. The plants are retarded in growth, the leaves are curled, bleached and accumulate anthocyanins at the rims. Photosynthesis is reduced and an analysis of metabolites shows a five to tenfold increase in leaf sucrose and starch content and an up to 100-fold increase in hexoses. A similar accumulation of soluble carbohydrates was found as well when petioles of potato leaves were cold-girdled. Exsudation experiments with excised leaves indicate a strong reduction in phloem transport. The reduced export strongly affects the supply of sink organs with sucrose, as the plants have a reduced root system and reduced tuber yield. The similarity to the phenotype of transgenic potato plants overexpressing a yeast invertase in the cell wall of leaves are striking.

To determine the apparent molecular mass of the sucrose transporter, an epitope tag has been fused to the C-terminus of the protein. This modification of the

carrier does not affect the transport activity. Western blot analysis using a primary antibody directed towards the epitope (c-myc) reveals an apparent molecular mass of the modified protein of about 50 kD (the tag contributes an additional 1.5 kD to the calculated molecular mass of 55 kD). Such an aberrant mobility in SDS-PAGE has been observed for several hydrophobic transport proteins. These results therefore do not contradict the biochemical evidence gathered in Poitiers, showing that the sucrose carrier appears as a 42 kD polypeptide on denaturating electrophoresis gels. Indeed, the addition of the tag may affect the hydrophobicity of the protein. A cDNA homologous to SoSUT1 was cloned, encoding a sucrose transporter from *Arabidopsis*. The encoded protein has been expressed and appears at 43 kD in SDS-PAGE.

II. The amino acid transporters of the plasma membrane

1. Physiology and biochemistry

The kinetics of valine uptake by sugar beet leaf discs were characterized gave evidence for two apparent saturable phases (K_{m1} = 0.21 mM; K_{m2} = 9.7 mM; V_{max1} and V_{max2} = 190 and 4500 pmol valine.min^{-1}.cm^{-2} leaf area respectively), superimposed by a diffusional phase.

Experiments with plasma membrane vesicles indicate that the amino acid/amino acid transporter complex possesses a positive charge at pH 5.5 but is neutral at pH 7.5. Because the amino acid remains zwitterionic within this pH range, it is assumed that the charge of the protein is modified between pH 5.5 and pH 7.5. These studies may be extended to precise the functional pK of the amino acid transporter.

2. Molecular biology

A PCR approach was employed to clone plant homologues of a family of amino acid permease found in fungi, bacteria and mammals (Rothamsted). Oligonucleotide mixtures corresponding to four conserved domains were synthesized and used in different combinations to amplify DNA fragments from a variety of plant templates. In preliminary experiments, a fragment of the proline uptake (*prn*B) gene of *Aspergillus* was cloned and used as a template to optimize the conditions for the PCR reaction. Although many fragments of appropriate sizes were amplified from the plant templates and sequenced, none had significant homology to amino acid permeases.

Oocytes from *Xenopus levis* have been used extensively for the heterologous expression of mammalian transport systems and have been an important tool for the characterization of transporters and, in some cases, for cloning the corresponding genes. However, at the time that this project was initiated, no plant transporter had been shown to be functionally expressed in oocytes. The STP1 mRNA, transcribed *in vitro* from the Arabidopsis hexose transporter STP1 cDNA, was injected into oocytes and successfully expressed. The finding that a plant membrane protein could be synthesized and targeted correctly to the oocyte plasma membrane to yield a functional transporter clearly demonstrated for the first time the utility of *Xenopus* oocytes as a heterologous expression system for the identification and characterization of plant transporters and as a potential tool for cloning their genes. However, using different batches of oocytes and different preparations of barley polyA$^+$ RNA, attempts to detect the expression of a plant amino acid carrier by monitoring uptake of radiolabelled amino acids into the injected cells were unsuccessful.

Complementation studies using a gap$^-$ (general amino acid permease) yeast strain have allowed the identification and characterization of several cDNA clones mediating amino acid transport, starting from an *Arabidopsis* cDNA library. Two of these clones (AAP1 and AAP2) have been fully characterized. There is at least one gene family that consists of more than five members in Arabidopsis. The genes seem to be differentially expressed and differ in their substrate specificity. One member seems to be a transporter with a broad specificity but transporting basic amino acids with only low affinity, whereas another member seems to be a general amino acid permease.

III. Ion transport at the plasma membrane

1. Physiology and mutagenesis

The approach developed in Milan allowed the isolation and the study of mutants of *Arabidopsis thaliana* altered in transport processes. This required a preliminary extensive characterization of ion transport in *Arabidopsis* to control the screening procedures.

Mutants putatively altered in K^+ uptake have been searched in M_2 progenies from EMS mutagenized M_1 plants by selecting braditroph seedlings in a medium containing NH_4^+ and sub-optimal K^+ concentrations or by selecting resistants to toxic concentrations of compounds such as Cs^+ or tetraethylammonium which interfere with K^+ uptake.

Selection for mutants putatively altered in the proton pump activity or in its sensitivity to the stimulating toxin FC has been performed by selection of resistants to toxic cations (paraquat) whose uptake is promoted by fusicoccin (FC). Up to now, two mutants (Cs 112 and 5-2) have been isolated with Cs+ and paraquat-FC selection.

Cs 112 is a monogenic recessive mutant, resistant to cesium when heterozygote ; when homozygote the mutant is sensitive to cesium and shows a pale-green phenotype with reduced growth which can be rescued by high level of K^+ (5 mM) in the medium. No further studies have been done for the moment on this mutant.

The 5-2 mutant is a monogenic mutant of *A. thaliana*, partially insensitive to FC, as far as resistance to FC-hygromycin, FC-TPP, FC-induced wilting, FC-induced cell expansion are concerned. The inheritance of the trait is dominant with the FC-hygromycin test and semidominant with the wilting test.

To explain the reduced response of 5-2 to FC, H^+ extrusion activity, K^+ influx and Em were measured in the mutant and in the wild type (*wt*). The results showed that the reduced response of 5-2 does not depend on some hindrance for FC to reach its site of action, or on a decreased affinity of the FC receptor for the toxin. Likewise, the reduced response of H^+ and K^+ transport to FC in the mutant does not depend on an impairment of the K^+ absorption system. The 5-2 mutation can influence the H^+ extrusion system independently of the presence of FC.

Specific activities of plasma membrane H^+-ATPase, Ca-ATPase and FC-binding were measured in plasma vesicles purified from 5.2 and *wt* mature leaves. While specific activities of Ca-ATPase and FC-binding were similar in the two genotypes, specific activity of plasma membrane H^+-ATPase was always about half in 5.2 compared to *wt*. The V_{max} of the enzyme in 5-2 is lower than in the wild type.

Other parameters (Km, pH dependence, sensitivity to vanadate, FC, proteolysis) were similar in 5-2 and in the wil type. This indicates that the 5-2 mutation determines a decrease (about 50%) in the amount of plasma membrane H^+-ATPase.

2. Molecular approaches on the nitrate transporter

Electrophysiological and uptake measurements made with oocytes injected with either barley root poly(A)$^+$ RNA or mRNA encoding the *Aspergillus* crnA nitrate transporter did not indicate the expression of a nitrate transporter. The Rothamsted group therefore turned towards the expression of the *Aspergillus* crnA gene in yeast.

PCR was used to obtain a DNA fragment containing the crnA coding region with minimal 5' and 3' untranslated regions. The resulting fragment was cloned into the yeast expression vector pYES2 (Invitrogen). Expression of the crnA gene was induced by growing the yeast cells in glucose-free medium containing D-galactose as the carbon source and several alternative procedures were then used to assay for functional expression. No functional expression was obtained.

To ensure that no mutations had inadvertently been introduced into the crnA coding region by the PCR approach, the fragment was sequenced. No point mutations were found, but when compared to the published partial crnA cDNA sequence, the sequence of the PCR fragment showed an apparent 2 bp deletion located 109 bp from the termination codon suggesting that the published sequence is incorrect. (This has since been confirmed by the authors). The 2 bp 'deletion' changes the reading frame of the C-terminus and predicts a protein product 24 amino acids larger than previously thought.

A search of the Swissprot database using the revised crnA protein sequence revealed several regions of homology with other membrane transporters. Alignment of these regions led to the identification of two motifs that were sufficiently conserved to allow the design primers for a PCR approach to cloning crnA homologues from barley. A 135 bp fragment was amplified from poly A$^+$ RNA extracted from nitrate-induced barley roots. Sequence analysis of the PCR fragment indicated homology (31% identity) to crnA, and the fragment hybridized to a small number of restriction fragments in a Southern blot of total barley DNA, confirming that it originated from the plant and was not a contaminant. The PCR fragment was used to screen a barley root cDNA library, and a number of hybridizing clones were identified. Preliminary results indicate that at least one of the clones (pBCH1) hybridizes to a nitrate-inducible mRNA species of about 2.2 kb. Sequencing of pBCH1 has shown it to contain an open reading frame that would encode a 544 residue hydrophobic polypeptide of 59 kD. Comparison of the predicted amino acid sequence with crnA shows the proteins to be 32.0% identical and 58.1% similar. From sequence analysis of other clones, it appears that at least three members of this gene family exist in barley and this is supported by the results of hybridization to Southern blots of barley genomic DNA. Initial attempts to obtain functional expression of pBCH1 in oocytes were unsuccessful. Based on these results, and those obtained with oocytes, we are now considering the possibility that additional factor(s) not present in the heterologous expression systems are required for the correct processing, targeting or activation of the crnA/BCH1 membrane protein.

IV. The malate transporter of the tonoplast

1. Malate transport in C3 plants

In Toulouse, vacuolar malate transport was analyzed by 'exchange diffusion' experiments. using tonoplast vesicles from *Catharanthus roseus*. These experiments showed the **reversibility** of the carrier and gave us the opportunity to test its specificity. Maleate, fumarate and succinate had no effect whereas they strongly inhibited malate influx into isolated vacuoles or tonoplast vesicles. Only malate, citrate and to a lesser extent quinate were able to mobilize the internal malate. A histidyl residue(s) is involved either in the binding or the translocation of malate, and the protonation of the histidyl residue is essential to provide a high rate of malate transport.

The kinetics of mutual inhibition between malate and citrate uptake, and the different sensitivity to protein-modifying reagents of either ^{14}C malate or ^{14}C citrate suggest that malate and citrate, which are the main accumulated organic acids in the *Catharanthus* vacuoles, cross the tonoplast by means of different carrier systems.

A photolyzable malate analogue, N-(4-azido-salicylyl) aspartic acid competitively inhibited the uptake of radioactive malate into purified tonoplast vesicles (K_i = 3.5 mM, compared to a Km of 2 mM for malate uptake). When iodinated, the malate analog was found to be still photolyzable and behaved as a competitive inhibitor of malate uptake (K_i = 1.4 mM). Photolysis of ^{125}I-labeled analog in the presence of purified tonoplast vesicles labeled several polypeptides separated by SDS-PAGE. However, only one polypeptide (40 kDa molecular weight) was totally protected from labeling by the inclusion of the substrate (malate) in the photolysis medium and also by the inclusion of 2 mM DEPC. The purification of the 40 kDa polypeptide, which may be the tonoplast malate carrier is underway.

2. Malate transport in CAM plants

Recent evidence suggested that malate influx into the vacuole of CAM plants (for example *Kalanchoë daigremontiana*, Crassulaceae) is mediated by a voltage-dependent, malate-selective ion channel. Since ion-channel proteins are normally present at extremely low levels, the strategy of the Oxford group has been to make use of molecular cloning techniques to attempt to identify the gene(s) encoding the vacuolar malate channel.

Monoclonal antibodies generated against the tonoplast of *K. daigremontiana* were tested for their ability to inhibit malate uptake across this membrane. Two complete rounds of monoclonal antibody production were performed. 5 positive cell lines inhibited malate transport after differential screening , but they ceased producing antibodies, possibly because the antibodies themselves were toxic to the hybridoma cells. One of the positive cell lines was successfully subcloned twice. Antibody from this cell line inhibited malate transport by 24%, the largest inhibitory effect observed for any of the cell-culture supernatants. This antibody reacted with a single polypeptide of 43 kDa in Western blots against tonoplast membrane. The limited amount of antibody available from this cell line will now be used to screen a cDNA expression library recently constructed from leaves *K. daigremontiana* (see below).

Patch-clamp studies in Oxford have shown that the tonoplast of leaf mesophyll cells of *Kalanchoë daigremontiana* possesses a voltage-dependent, malate-selective ion channel. The channel is active only at inside-positive vacuolar membrane

potentials and shows a high selectivity for four-carbon, trans 1,4-dicarboxylates. Since the properties of the vacuolar malate channel exactly match those predicted from earlier studies of the anion-dependence of ATP- or PP_i-driven H^+ influx into tonoplast vesicles, we believe that this channel represents the principal pathway for malate transport from the cytosol into the vacuole.

To determine whether expression of the vacuolar malate channel might be developmentally regulated, malate currents were measured in the whole-vacuole configuration in vacuoles from mature leaves of *K. daigremontiana* performing CAM and compared with those from young leaves, which perform C_3 photosynthesis. Although the total vacuolar malate current was greater in vacuoles from mature leaves, this turned out to be purely a function of vacuole size : expressed on a unit-area basis, the current flux density was no greater than that across the tonoplast from young leaves. Therefore, differential screening of gene transcripts from leaves at different stages of development is not likely to be a useful strategy for identifying the gene(s) encoding the vacuolar malate channel in *Kalanchoë*.

As part of the screening procedure for selecting monoclonal antibodies (above), a membrane-filtration assay was developed for measuring the rate of passive (non-energized) malate uptake by tonoplast vesicles from *Kalanchoë daigremontiana*, and the characteristics of malate (= ^{14}C] succinate) uptake by these vesicles were studied (Bettey and Smith, 1993). Inhibition by pyridoxal phosphate showed pseudo-first-order kinetics with a reaction order of 1.03 and was completely protectable by malate (with a Kd for substrate protection very close to the K_m for substrate transport), indicating the involvement of a least one critical lysine residue in the anion-recognition site of the transport protein.

Many of the properties of malate uptake by tonoplast vesicles from *Kalanchoë* agree with observations made by the Toulouse group with tonoplast preparations from the C_3 plant *Catharanthus roseus*. In C_3 plants, the malate-uptake system at the tonoplast has usually been considered to represent some form of 'carrier' or 'permease' capable of working in an exchange-diffusion mode (see above). This has prompted us to consider whether the tonoplast of CAM plants possesses a unique malate-selective ion channel, or whether both types of transport system reside in the same membrane.

The tonoplast of CAM plants also exhibits the carrier-type transport system. Indeed, in membrane-filtration assays, rates of malate uptake by *Kalanchoë* tonoplast vesicles increase with decreasing external pH, as found for *Catharanthus* tonoplast vesicles; in contrast, recent patch-clamp measurements with *Kalanchoë* vacuoles have shown that whole-vacuole malate currents decrease as external ('cytosolic') pH decreases below pH 7.5.

Altogether, the data suggest that two different malate-transport systems reside in the *Kalanchoë* tonoplast: a voltage-dependent malate channel mediating influx, and a carrier system, active at low pH (pH 5.5), mediating malate **efflux** from the vacuole. This hypothesis will be further tested by a collaboration between the Toulouse and Oxford groups.

MAJOR SCIENTIFIC BREAKTHROUGHS

The main objectives of the programm have been achieved. Considerable progress has been made in the biochemistry of several transporters (sucrose transporter of the plasma membrane, malate transporter of the tonoplast). The cloning of several transporters (sucrose, amino acids) opens wide avenues for future research concer-

ning site-directed mutagenesis, transgenesis and expression studies. This is specially important because of the role played by these proteins in the morphogenesis and the productivity of the plant. Progress made in biochemical, molecular biology and mutagenesis approaches also allows to expect in the near future the cloning of other transporters, or of regulatory proteins controlling the transport of ions and organic solutes.

MAJOR COOPERATIVE LINKS

During the programm, all participants interacted closely during several general meetings at about 6 months intervals. Many ideas, techniques (phase partition, free-flow electrophoresis, two-dimensional gel electrophoresis, molecular biology), and tools (antibodies, molecular probes, transgenic plants, plasmids, yeasts) were shared and exchanged, especially between Poitiers and Berlin. Many new tools and concepts will be still exploited in the near future (library of monoclonal antibodies against tonoplast proteins, cDNA library from *Kalanchoë*, microsequencing of particular membrane proteins identified in the tonoplast or in the plasma membrane during the course of this programm, *Arabidopsis* mutants, etc...). All participants have agreed to extend their collaboration as far as possible in the future.

PUBLICATIONS*

* We have only listed the papers where the support of the work by the present programm has been explicitely stated

Joint publications

Frommer, W.B., Hummel, S., Lemoine, R. and Delrot, S. (1994). Developmental changes in the two dimensional protein pattern of plasma membrane vesicles between sink and source leaves from sugarbeet. Plant Physiol. Biochem. (in press).

Lemoine, R., Gallet, O., Gaillard, C., Frommer, W. and Delrot, S. 1992. Plasma membrane vesicles from source and sink leaves: changes in solute transport and polypeptide composition. Plant Physiol. **100**, 1150-1156.

Individual publications

Beffagna, N. and Romani, G. (1993) *Arabidopsis* suitability for studies on intracellular pH regulation: correlation between H^+ pump activity, cytosolic pH and malate level in wild type and in a mutant partially insensitive to fusicoccin. Plant, Cell and Environ. (in press).

Beffagna N., Romani, G. and Lovadina, S. (1993) Inhibition of malate synthesis by vanadate: implications for the cytosolic alkalinization induced by vanadate concentrations partially inhibiting the plasmalemma H^+ pump. J. Exp. Bot. **44**, 1535-1542.

Boorer, K.J., Forde, B.G., Leigh, R.A. and Miller, A.J. (1992) Functional expression of a plant plasma membrane transporter in *Xenopus* oocytes. FEBS Lett. **302**, 166-166.

Cerana, R. and Colombo, R. (1992) K^+ and Cl^- conductance of *Arabidopsis thaliana* PLASMA MEMBRANE at depolarized voltages. Bot. Acta **105**, 273-277.

Colombo, R., Cerana, R. and Bagni, N. (1992) Evidence for polyamine channels in protoplasts and vacuoles of *Arabidopsis thaliana* cells. Biochem. Biophys. Res. Comm. **182**, 1187-1192.

Colombo, R. and Cerana R. (1991) Inward rectifying K^+ channels in the plasma membrane of *Arabidopsis thaliana*. Plant Physiol. **97**, 1130-1135.

De Michelis M.I., Carnelli A. and Rasi-Caldogno F. (1993) The Ca^{2+} pump of the plasma membrane of *Arabidopsis thaliana* : characteristics and sensitivity to fluorescein derivatives. Bot. Acta **106**, 20-25.

Dietz, K-J, Canut, H., and Marigo, G. (1992) Identification of an essentiel histidine residue at the active site of the tonoplast malate carrier in cells. J. Memb. Biol. **129**, 137-143.

Frommer ,W.B., Hummel, S. and Riesmeier, J. (1993) Yeast expression cloning of a cDNA encoding a broad specificity amino-acid permease from *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **90**, 5944-5948.

Gallet, O., Lemoine, R., Gaillard, C., Larsson, C., and Delrot, S. (1992). Selective inhibition of active uptake of sucrose into plasma membrane vesicles by polyclonal sera directed against a 42 kD plasma membrane polypeptide. Plant Physiol. **98**, 17-23

Gomarasca, S., Vannini, C., Venegoni, A., Talarico, A., Marrè, M.T. and Soave, C. (1993) A mutant of *Arabidopsis thaliana* with a reduced response to fusicoccin. I. Plant Physiol. **103**: 165-170.
Laloi, M, Delrot, S. et M'Batchi, B. 1993. Characterization of sugar efflux from sugar beet leaf plasma membrane vesicles. Plant Physiol. Biochem. **31**, 731-741.
Li, Z-S., Gallet, O., Gaillard, C., Lemoine, R. andt Delrot, S. 1992. The sucrose carrier of the plant plasmalemma. Partial purification and reconstitution into proteoliposomes. Biochim. Biophys. Acta **1103**, 259-266.
Olivari, C., Pugliarello, M.C., Rasi-Caldogno, F. and De Michelis, M.I. (1993) Characteristics and regulatory properties of the H^+-ATPase in a plasma membrane fraction purified from *Arabidopsis thaliana*. Bot. Acta **106**, 13-19.
Rasi-Caldogno, F. and De Michelis, M.I. (1992) The plasma membrane Ca^{2+} pump: biochemical characteristics and regulatory properties. In 'Biochemical Mechanism Involved in growth Regulation'. D. Chiatante, J. Gallon, C. Smith and G. Zocchi eds., Phytochemical Society of Europe, Oxford University Press, (in press).
Riesmeier, J., Willmitzer, L. and Frommer, W.B. (1992) Isolation and characterization of a sucrose carrier cDNA from spinach by functional expression in yeast. EMBO J. **11**, 4705-4713.
Riesmeier, J.W., Hirner, B. and Frommer, W.B. (1993) Expression of the sucrose transporter indicates a role in phloem loading. Plant Cell, (in press).
Riesmeier, J.W., Willmitzer, L. & Frommer, W.B. (1994) Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning. EMBO J., (in press).
Sakr, S., Lemoine, R., Gaillard, C. and Delrot, S. (1993). Effect of cutting on solute uptake by plasma membrane vesicles from sugar beet (*Beta vulgaris* L.) leaves. Plant Physiol. **103**, 49-58.

AREA C:

CELLULAR BIOLOGY

BIOTECHNOLOGY OF ANIMAL CELLS

Construction of artificial chromosomes for higher eucaryotic cells (BIOT CT-910259)

COORDINATORS:
H.J. LIPPS, Univ. of Tübingen, Tübingen, D
P. DONINI, Univ. of Rome, Rome, I
PARTICIPANTS:
F. AZORIN, CSIC, Barcelona, ES
H.J. COOKE, MRC, Edinburgh, UK
F. GRUMMT, Univ. of Würzburg, Würzburg, D
M. SHIRLEY, AFRC, Pirbright, UK

BACKGROUND INFORMATION

No satisfactorily autonomously replicating vectors exist for higher eucaryotic cells and organisms and it may well be that the ideal eucaryotic vector may have to mimic a chromosome. The minimal requirements for an artificial chromosome are sequences to allow and regulate its autonomous replication, centromeric sequences to guarantee the mitotic and meiotic stability of the chromosome, and telomeric sequences needed for its replication as a linear DNA molecules.

OBJECTIVES AND PRIMARY APPROACHES

The general objective of this proposal was the construction of linear vectors and artificial chromosomes for mammalian cells by assembling biological relevant DNA sequences from chromosomes and naturally occuring linear plasmids. Unlike conventional circular eucaryotic vectors that integrate in an unpredictable and random fashion into the host genome, an artificial chromosome would replicate extrachromosomally as an independent functional unit. Therefore, these constructs will allow directed and reproducible engineering of higher eucaryotes, will be useful in the correct expression of medically important proteins, will help to understand differentiation processes and will provide a safe vector system for the genetic therapy of some human diseases. In addition, basic knowledge about the eucaryotic chromosome was acquired in the course of this project.

RESULTS AND DISCUSSION

Initially major efforts were made towards a structural and functional analysis of the functional chromosomal elements, **telomeres,** replication origins and centromeres. In a following step these elements were assembled into linear constructs and tested for physical stability and replication in various organisms. Further work concentrated on structural studies of defined DNA sequences and on the construction of novel eucaryotic circular expression vectors.

Of the three functional chromosomal elements telomeres, needed for correct replication of chromosomal ends and for the protection of chromosomes from exonucleolytic digestion, are the best characterized. They are isolated from many organisms and consist of tandemly repeated short sequences which are highly conserved through evolution, the consensus sequence being Tm(A)Gn. Most of the work in the past years concentrated on telomere sequence requirements for replication and stability and on the analysis of telomeric DNA structure adopted *in vitro* and *in vivo*. For this analysis a high copy number linear yeast vector constructed in the course of this project proved most useful. In addition, the possible effects of proximity of a telomere on the expression of genes incorporated into a

mammalian artificial chromosome was investigated. **Centromeres** are still poorly understood and so far no functional centromeric sequence from higher eucaryotes have been isolated. However, techniques to handle such large structures are now available which will allow the isolation and characterization of centromeres. Although many putative **origins of bidirectional replication** have been mapped in several organisms, so far no functional origin of replication has been isolated from mammalian cells. However, there is increasing indirect evidence that the amplification promoting sequences originally isolated from the non-transcribed spacer region of murine rDNA act *in vivo* as origin of bidirectional replication. Other approaches to overcome the problem of a functional origin of replication was the use of viral origins of replication from SV40 or BPV-1 and, under the assumption that a functional origin of replication occurs every few hundred kilobases in chromosomal DNA, the insertion of large mammalian DNA fragments into YAC vectors.

A number of linear constructs have been constructed and tested in various biological systems. These constructs are:

1. **Viral based linear constructs:** SV40 or BPV-based vectors were linearized and human telomeric sequences ligated to the termini.

2. **Constructs based on putative mammalian origins of replication:** The viral origin of replication of a BPV-based vector was removed and replaced by the amplification promoting sequences.

3. **YAC-based linear constructs:** YAC-vectors were constructed with removable yeast telomeres capping vertebrate telomeric sequences. These vectors can be propagated in yeast cells and uncapped into a form functional in mammalian cells. From a number of these constructs it could be shown that they are physically stable and even replicate as an episomal linear vector in mammalian cells.

Further work concentrated on the structural analysis of DNA sequences involved in recombination and on the construction of novel eucaryotic expression vectors. Using an SV40 based approach the effect of simple repeated DNA sequences on homologous recombination was analyzed. It could be shown that the frequency of homologous recombination is increased by an order of magnitude in the presence of repeating d(GA.TC)n sequences. Based on the observation that BPV-1 replicates extrachromosomally in *Xenopus* embryos, an expression vector allowing the identification of genes involved in early embryogenesis of *Xenopus* was constructed. It could be demonstrated that the amplification promoting sequences promote plasmid amplification under non selective conditions in mammalian cells. The amount of transcript expressed from such a vector is directly proportional to its copy number. A vector containing part of the HIV-1 genome in antisense was constructed. HIV-1 replication in a human T-lymphocyte cell line was nearly abolished when antisense RNA was expressed from such a high copy number vector.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

All the results obtained contribute significantly to our understanding of chromosome structure and function and therefore are of major scientific interest. Moreover, our group was able to assemble several types of linear constructs which are presently being tested in various biological systems. This type of linear vector may prove useful for directed genetic engineering of higher eucaryotic cells, may improve the efficiency in creating transgenic animals and may be the ideal vector for safe and efficient genetic therapy of some human diseases, and as such will be

of great interest for industry and medicine. In addition, in the course of this project important questions concerning recombination processes and site specific integration were answered. The novel circular expression vectors constructed are already now used for expression of heterologous proteins in mammalian cells and in gene therapy projects.

MAJOR COOPERATIVE LINKS

There was strong interaction between all participating groups as well as between other groups of the T-project 'Animal cell biotechnology'. A regular exchange of material and staff took place and in addition to the meetings within the framework of the T-project, our group met at least twice a year. It is obvious, that this type of research can only be performed in the framework of an EC-project and the experience made in the course of the last years showed that this type of interaction proved not only useful for our group, but also for other groups funded by the EC, for european science and for european industry.

PUBLICATIONS

Joint publications

Huertas, D., Lipps, H.J. and Azorin, F.: Characterization of the structural conformation adopted by (TTAGGG)n telomeric DNA repeats of different length in closed circular DNA. J. Biomolecular Structure and Dynamics, (in press).
Meyer, J., Nick, S., Stamminger, T., Grummt, F., Jahn, G. and Lipps, H.J. (1993): Inhibition of HIV-1 replication by a high copy number vector expressing antisense RNA for reverse transcriptase. Gene **129**, 263-268.
Schmid, M., Steinbeisser, H., Ascenzioni, F., Trendelenburg, M.F. and Lipps, H.J. (1991): Native yeast telomeres are sufficient to stabilize linear DNA in Xenopus oocytes. Gene **106**, 121-124.
Weber, S., Meyer, J., Schmid, M., Cooke, H. and Lipps, H.J. (1993): A linear vector carrying human telomeric sequences replicates extrachromosomallyin Xenopus eggs. Cell Biol. Int. Rep. **17**, 623-624.

Individual publications

Abken, H., Weidle, U.H. and Grummt, F.: Tumorigenic transformation of established murine fibroblasts by transfection with amplification promoting DNA elements, (submitted).
Ascenzioni, F., Pucci, L., Garulli, B., Guerrini, A.M. and Donini, P.: Replication of linear SV40-based plasmids in monkey cells, (submitted).
Beltran, R., Martinez-Balbas, A. Bernues, J., Bowater, R. and Azorin, F. (1993): Characterization of the zinc-induced structural transition to*H-DNA at a $d(GA\text{-}CT)_{22}$ sequence. J. Mol. Biol. **230**, 966-978.
Bernues, J., Beltran, R. and Azorin, F. (1991): SV40 recombinants carrying a $d(CT.GA)_{22}$ sequence show increased genomic instability. Gene **108**, 269-274.
Carrera, P., Martinez-Balbas, Portugal, J. and Azorin, F. (1991): Identification of sequence elements contributing to the intrinsic curvature of the mouse satellite DNA repeat. Nucl. Acids Res. **19**, 5639-5644.
Carrera, P. and Azorin, F.: Structural characterization of intrinsically curved AT-rich DNA sequences. J. Biol. Chem., (in press).
Casasnovas, J.M., Huertas, D., Ortiz-Lombardia, M., Kypr, J. and Azorin, F. (1993): Structural polymorphism of d(GA.CT)n sequences: intramolecularand intermolecular association of the individual strands. J. Mol. Biol. **233**, 671-681.
Casanovas, J.M. and Azorin, F. (1991): The effect of simple repeating d(CG.GC)n, d(GA.TG)n and d(A.T)n DNA sequences on the nucleosomal organisation of SV40 minichromosomes. DNA and Cell Biol. **10**, 751-756.
Cooke, H.J.(1992): Using chromosome features in genome mapping. TIBTECH **10** ,23-27.
Euntari, F. and Donini, P.: *In vitro* identification of a protein of *Saccharomyces cerevisiae* that interacts specifically with the G-rich strand of the telomere. Res. Microbiol., (in press).

Farr, C.J., Stevanovic, M., Thomson, E.J., Goodfellow, P.N. and Cooke, H.J. (1992): Telomere-associated chromosome fragmentation: applications in genome manipulation and analysis. Nature Genetics **2**, 275-282.
Ferrer, N., Azorin, F., Villasante, A., Gutierrez, C. and Abad, J.P.: Centromeric dodecasatellite DNA sequences from fold-back structures. EMBO J., (in press).
Gögel, E., Müller, C., Kunkel, E. and Grummt, F.: Mapping an origin of bidirectional replication in mouse rDNA repeat units, (submitted).
Guerrini, A.M., Campoeneschi, B., Ascenzioni, F., Picolella, E. and Donini, P. (1993): Subtelomeric as well as telomeric sequences are lost from chromosomes in proliferating B-lymphocytes. Hum. Mol. Gen. **2**, 455-460.
Guerrini, A.M., Ascenzioni, F., Gallo, T. and Donini, P. (1993: Determination of plasmid copy number in yeast transformants by means of agarose plugs. Nucl. Acids Res. **25**, 6047-6048.
Guerrini, A.M., Sscenzioni, F., Gallo, T. and Donini, P.: Linear plasmids with synthetic human telomeres replicate autonomously in Saccharomyces cerevisiae, (submitted).
Hemann, C., Gärtner, E., Weidle, U.H. and Grummt, F. (1994): A high copy expression vector based on amplification promoting sequences. DNA and Cell Biol. **13**, 437-449.
Huertas, D., Bellsolell, L., Casanovas, J.M., Coll, M. and Azorin, F. (1993): Short d(GA)n DNA sequences form antiparallel stranded homoduplexes stabilized by the formation of GA pairs. EMBO J. **12**, 4029-4038.
Kipling, D. Ackford, H.E., Taylor, B.A. and Cooke, H.J. (1991) Mouse minor satellite DNA genetiocally maps to the centromere and is physically linked to the proximal telomere. Genomics **11**, 235-241.
Kipling, D., Wiloson, H., Mitchell, A.R., Taylor, B. and Cooke, H.J. (1994): Mouse centromere mapping using oligonucleotide probes that detectvariants of the minor satellite. Chromosoma **103**, 46-55.
Koehler, U., Abken, H., Grummt, F. and Weidle, U.H.: Amplification and promoting sequences can mediate the transposition of plasmids located in homogenerously staining region into a plethora of new chromosomal integration sites. Cytogenetics and Cell Genetics, (in press).
Maercker, C. and Lipps, H.J. (1994): Analysis of the subtelomeric regions of the macronuclear gene-sized DNA-molecules from the hypotrichous ciliate Stylonychia lemnae. Implications for the DNA fragmentation process during macronuclear development? Dev. Genetics **14**, 378-384.
Martinez,-Balbas, A. and Azorin, F. (1993) The effect of zinc on the secondary structure of d(GA.TC)n DNA sequences of different length. A model for the formation of *H-DNA. Nucl. Acids Res. **21**, 2557-2562.
McGill, N.I.. Fantes, J. and Cooke, H.J. (1992): λCM8, a human sequence with putative centromeric function, does not map to the centromere but is present in one to two copies at qtr. Hum. Mol. Genet. **1**, 749-751.
McKay, S.J. and Cooke, H.J. (1992) hnRNP A2/B1 binds specifically to single stranded vertebrate telomeric repeat TTAGGG. Nucl. Acids Res. **20**, 6461-6464.
McKay, S.J. , Thomson, E. and Cooke: Sequence homologies and linkage group conservation of the human mouse CENPC genes. Genomics, (in press).
Müller, F. and Grummt, F. (1991): Mapping eukaryotic replication origins *in vivo* by size analysis of purified nascent DNA strands. DNA and Cell Biology **10**, 149-157.
Pesch, J., Gärtner, E. and Grummt, F.: p5 interacts with an amplification promoting element from the murine rDNA-locus, (submitted).
Schmid, M., Steinbeisser, H., Epperlein, H.H., Trendelenburg, M.F. and Lipps, H.J. (1992): An expression vector inhibits gene expression in Xenopus embryos by antisense RNA. Roux Arch. Dev. Biol. **201**, 340-345.
Sock, E., Wegner, M., Fortunato, E.A. and Grummt, F. (1993): Large T-antigen and sequences within the regulatory region of IC virus both contribute to the features of JC virus DNA replication. Virology **197**, 537-548.
Stoll, S., Zirlik, T., Maercker, C. and Lipps, H.J. (1993): The organization of internal telomeric repeats in the polytene chromosomes of the hypotrichous ciliate Stylonychia lemnae. Nucl. Acids Res. **21**, 1783-1788.
Stolzenburg, F., Dinkl, E. and Grummt, F. (1992): Nucleotide sequence of a mouse cDNA encoding the non-histone chromosomal high mobility group protein 2 (HMG-2). Nucl. Acids Res. **20**, 4927.

Stolzenburg, F., Gerwig, R., Dinkl, E. and Grummt, F.: Structural homologies and functional similarities between mammalian origins of replication and amplification promoting sequences. Chromosoma, (in press).

PATENTS

Grummt, F. and Weidle, U. (1992): Murine autonomously replicating sequences and vectors containing. Pat. No. 5,081,018 Washington DC, USA.
Grummt, F. and Weidle, U. (1993): Expressions vektor und Verfahren zur Expression von heterologen Proteinen in Säugertierzellen. Europ. Patent Nr. 306 648 B1.
Lipps, H.J., Grummt, F. und Meyer, J.: Amplifizierbarer Vektor gegen HIV-Replikation. Europ. Patentanmeldung Nr. P42 36 582.1.

Improved techniques for establishing a high expression production system for recombinant proteins from animal cells (BIOT CT-900185)

COORDINATOR:
H. HAUSER, Biotechnologische Forschung, Braunschweig, D
PARTICIPANTS:
M.J.T. CARRONDO, Inst. Biologia Experimental e Tecnologica, Oeiras, P
B. GRIFFITHS, Microbiology and Research, Salisbury GB
M. MISTLER, (Subc.), Schott Glaswerke, Mainz, D
U.H. WEIDLE, Boehringer Mannheim, Penzberg, D

BACKGROUND INFORMATION

A major bottleneck in the production of proteins for pharmaceutical use in mammalian cells is the productivity and stability of the used cells during fermentation at the technical scale. To ensure an economically acceptable process an integrated approach including genetic and cell culture technology is required.

OBJECTIVES AND PRIMARY APPROACHES

Evaluation of the relationship of genetic and technological parameters for the production of recombinant gene products in mammalian cell lines. Parameters: Chromosomal localization of the transfected genes and the methods of cultivation for long-term stability, strength of gene expression, product accumulation and cell physiology. The goal is a protocol for the construction and cultivation of cell lines with which to achieve optimal conditions for recombinant protein production at the technical scale.

Genetic works

Identification and targeting hyperactive chromosomal sites in SP2/O and BHK-21 cells with model genes into which the genes of interest could be integrated by homologous recombination.

- Construction of efficient expression vectors for the expression of antibody genes susceptible for gene targeting.
- Establishment of screening methods which allow the isolation of cell clones with stable and high expression of model genes.
- Establishment methods for targeting the genes of interest into the identified by recombination processes.
- Alternatively, creation of hyperactive chromosomal sites by making use of 'Scaffold-Attached-Regionelements (SARs).

Cell Technology

Evaluation of novel genetic constructs and cell lines.

- Evaluation of cell line in different biorector systems (Stirred tank bioreactors, spinner systems of single cell suspension, suspended aggregates and attached to microcarriers.
- Investigation of the relationship between physiological parameters and cell productivity.

— Analysis of the behaviour of recombinant cell lines immobilized in high density culture systems based upon porous carriers.

RESULTS AND DISCUSSION

Genetics

The strategy to detect hyperactive chromosomal regions was to introduce a single copy reporter gene into a large number of cell clones by random integration, to isolate those clones which show the highest expression of the reporter gene and to isolate the flanking regions of the integrated reporter gene. Cells were infected with a recombinant retrovirus expressing a secreted alkaline phosphatase gene (SEAP) and the puromycin resistance gene from a bicistronic expression unit. The screening procedure includes high concentrations of puromycin followed by an overlay assay which detects SEAP-overexpressing cell clones. Cell clones showing a high level of SEAP expression were isolated. The same procedure was carried out with NIH 3T3 and BHK-21 cells. Since BHK-21 cells are refractory to infection with murine retroviruses, these cells manipulated to express the ecotropic murine retrovirus receptor were chosen for these experiments. Stability of expression from all these SEAP overexpressing cell clones were tested for at least three months. A reduction of expression down to 30% of the initial value was observed. After evaluation of the cell clones by the partners in the project the chromosomal flanks of the cell clones were amplified by the inverted PCR procedure and cloned into cloning/sequencing vectors. The length of the inserts (5' and 3' boxes) were between 350 and 2100 bp. The intention was to use these flanking sequences for homologous recombination. Homologous recombination, as a method of targeting a gene into selected and characterized loci was first examined with NIH3T3 cells into a locus with an established positive influence on the expression of an integrated gene. Cloned DNA sequences adjacent to the vector had lengths of 700 bp and 1400 bp, respectively. After selection using the 'positive-negative' method 1 from 10-5 cells survived. These were pooled and tested for the event of homologous recombination by an optimized PCR method. In our experiments no positive clone was found.

Construction and properties of artificial chromatin domains: A variety of techniques has been used for the isolation of functional domain borders or SAR elements, resp. Various sets of these were applied for the construction of artificial chromatin minidomains. All combinations supported the stable expression of test genes under the control of different natural and an artificial promoters. SAR-specific effects depended on the strength of the SAR-scaffold interaction and sets of strong elements permitted the stable expression over extended periods of time (6 month). Stimulatory functions of butyrate depending both on SAR-strength and upon nature of the promoter were observed. The optimum combination of elements comprises constructs based upon the metallothionein I-promoter embedded in a minidomain with strong SAR-elements. Under these conditions, the isolated effects of SARs (6 fold enhancement) and butyrate (80 fold stimulation) combine in a synergistic fashion to a considerable increase of transcriptional rates. This system enables the combination of a growth period at moderate levels of expression with the potential of 1-5 mM concentrations of butyrate of boosting expression

Cell Technology

An evaluation of cell lines was carried out by relating physiological and physical factors in product expression to

(i) different cell lines containing the same genetic construct,
(ii) different genetic constructs in the same cell line (BHK-21, SP2/0) and
(iii) different culture systems.

Cells were grown in either free suspension, natural aggregate suspension, or on solid microcarriers. A comparison of aggregates with microcarrier cultures show that the production rate is higher in a controlled aggregate culture. Investigated products were recombinant IgG and SEAP. Suspended aggregated cultures have shown a superior scale up potential than microcarrier cultures as the range of power inputs for cell growth without damage is larger.

Depending on the growth conditions, the morphology of BHK cells changes significantly and this change in morphology seems to correlate with the specific production rate.

Significant differences found between cell clones derived from one transfection experiment indicate the importance of the chromosomal neighbourhood.

The relationship of environmental parameters of a fixed-bed bioreactor (FBR) system to the productivity of recombinant cell lines was investigated. Different types of Siran beads were used, non derivatized and derivatized. The effects of dilution and circulation rates, bead and inoculum sizes, and feed glutamine concentration were investigated. Increases in productivity seem to be associated with smaller beads and higher circulation rates. Fluidized bed bioreactor systems show a higher specific growth rate but a lower value of the production rate. Immobilization within a FBR system using Siran beads increases the production rate. The microenvironment within high density culture systems is more favourable for high product expression levels than microcarriers. This is irrespective of whether a high density system is used.

The use of SP2/0 cells was evaluated by using transfectants expressing SEAP. A variability in both expression of SEAP and a range of cultural characteristics has been analyzed with respect to the different cell clones derived from a single transfection experiment. In SP2/0 cells a high growth rate and high product expression rate are apparently associated with a high rate of glutamine metabolism. Examination of production rate and cell specific metabolic rate data suggested an association between high rates of SEAP production and glutamine catabolism. High growth rate and SEAP production rates are both associated with high values for glutamine uptake and production of glutamate, alanine and ammonium.

Investigation of interrelationship of genetic, physiological and physical factors in the regulation of high product expression shows that the choice of growth mode is important. It can have a profound effect upon both the relationship between specific growth rate and production rate.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The studies undertaken in this project underscore the strong influence of the chromosomal locus in which the recombinant gene has been integrated. SAR elements are not only hallmarks of high expression, but also can be used to create hyperactive artificial chromosomal domains.

The choice of growth mode can have a profound effect upon both the relationship between specific growth rate and production rate as well as the maximum value of the production rate. Particularly, cells grown as aggregates, have a markedly higher production rate as cells grown on microcarriers. This has obvious consequences for the development of commercial production processes. Long-term

studies also showed that the fixed-bed reactor system improved the stability of recombinant protein production.

MAJOR COOPERATIVE LINKS

More than five work exchanges by members of the participating groups in partner laboratories took place.

PUBLICATIONS

Joint publications

Moreira, J. L., Wirth, M., Fitzek, M. and Hauser, H.: Evaluation of reporter genes in mammalian cell lines. Methods Mol. Cell. Biol. **3**, 23-29 (1992)

Racher, A. J., Moreira, J. L., Alves, P. M., Wirth, M., Weidle, U. H., Hauser, H., Carrondo, M. J. T. and Griffiths, J. B.: Expression of recombinant antibody and secreted alkaline phosphatase im mammalian cells. Influence of cell line and culture system upon production kinetics. Applied Microbiology and Biotechnology **40**, *851-856 (1994)*

Individual publications

Alves, P. M., Moreira, J. L., Feliciano, A. S., and Carrondo, M. J. T.: A comparison of animal cell growth using microcarriers and suspended natural aggregates. In: R. E. Spier, J. B. Griffiths, W. Berthold (eds.), Animal Cell Technology: Products of Today, Prospects of Tomorrow, Butterworth & Co., 321-323 (1994).

Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Mielke, C., and Kohwi-Shigematsu, T.: Biological Significance of Unwinding Capability of Nuclear Matrix-Associating DNAs. Science **255**, 195-197 (1992)

Dietz, A., Bode, J., and Landsmann, J.: A Scaffold Attachment Region is associated with a T-DNA integration site in transgenic petunia. Nucleic Acids Res., (submitted).

Dirks, W., Wirth, M. and Hauser, H.: Bicistronic transcription units for gene expression in mammalian cells. Gene **128** 247-249 (1993)

Dirks, W., and Hauser, H.: Equimolar expression of two protein chains in mammalian cells. In: R. Spier, J. B. Griffiths, and W. Berthold (eds.), Animal Cell Biotechnology: Products of Today, Prospects for Tomorrow, (1994), Butterworths, Oxford, 610-616

Dirks, W., Schaper, F., and Hauser, H.: A new hybrid promoter directs identical mRNA starts in mammalian cells and in vitro. Gene, (in press).

Dirks, W., Schaper, F., Kirchhoff, S., Morelle, C., Gross, G., and Hauser, H.: A multifunctional vector family for gene expression in mammalian cells. Gene, (in press).

Dirks, W., Mielke, C., Karreman, S., Haase, B., Wirth, M., Lindenmaier, W., and Hauser, H.: Applications of expression vectors containing bicistronic transcription units in mammalian cells. In: Schering Research Foundation Workshop, N. E. Fusenig, and H. Graf (eds.), Springer-Verlag, Heidelberg, (in press).

Kay, V., and Bode, J.: Binding specificities of a nuclear scaffold: Supercoiled, single stranded and SAR-DNA. Biochemistry, (in press).

Klehr, D., Schlake, T., Maaß, K., and Bode, J.: Scaffold-Attached Regions (SAR elements) mediate transcriptional effects due to butyrate. Biochemistry **31**, 3222-3229 (1992)

Moreira, J. L., Alves, P. M., Aunins, J. G., and Carrondo, M. J. T.: Aggregate suspension cultures of BHK cells. In: R. E. Spier, J. B. Griffiths, G. MacDonald (eds.), Animal Cell Technology: Developments, Processes and Products, Butterworth & Co., 411-413 (1992)

Moreira, J. L., Alves, P. M., Aunins, J. G., and Carrondo, M. J. T.: Baby hamster kidney aggregation in fermentation. In: F. Vardar-Sukan, S. Suha-Sukan (eds.), Recent Advances in Industrial Applications of Biotechnology, Kluwer Academic Publishers, NATO-ASI series, Serie E, Vol. **210**, 537-538 (1992)

Moreira, J. L., Feliciano, A. S., Alves, P. M., Aunins, J. G., and Carrondo, M. J. T.: Hydrodynamic control of suspended natural animal cell aggregates. In: R. E. Spier, J. B. Griffiths, W. Berthold (eds) Animal Cell Technology: Products of Today, Prospects of Tomorrow, Butterworth & Co., 397-401 (1994)

Moreira, J. L., Feliciano, A. S., Aunins, J. G., and Carrondo, M. J. T.: Long term cultures of BHK suspended aggregates, in R. E. Spier, J. B. Griffiths, W. Berthold (eds) Animal Cell Technology: Products of Today, Prospects of Tomorrow, Butterworth & Co., 507-510 (1994).

Moreira, J. L., Alves, P. M., Aunins, J. G., and Carrondo, M. J. T.: Changes in animal cell natural aggregates in suspended batch cultures, Appl. Microb. Biotechnol., (in press).
Moreira, J. L., Alves, P. M., Rodrigues, J. M., Cruz, P. E., and Carrondo, M. J. T.: Studies of baby hamster kidney (BHK) natural cell aggregation in suspended batch cultures, Annals N. Y. Acad. Sci., (in press).
Racher, A. J., and Griffiths, J. B.: Investigations of parameters affecting a fixed bed bioreactor process for recombinant cell lines. Cytotechnology **13**, 125-131 (1993).
Racher, A. J., and Griffiths, J. B.: Development of a bioreactor process for recombinant cell lines. In: Animal Cell Technology, Products of Today, Prospects for Tomorrow, (eds.) R. E. Spier, J. B. Griffiths, W. Berthold, Butterworth, Oxford, 248-251 (1994)
Racher, A. J., and Griffiths, J. B.: Recombinant antibody expression in a fixed bed perfusion system. In: Animal Cell Biotechnology: Basic and applied Aspects, K. Okumura (ed.), Kluwer Academic Publishers, Vol. 6, (in press).
Rzepecki, R., Smidzinski, R., Bode, J., and Szopa, J.: The 65 kDa protein affected endonuclease tightly associated with plant nuclear matrix. J. Plant Physiol. **139**, 284-288 (1992)
Szopa, J., Bode, J., Kay, V., Kozubek, A., and Sikorski, A. F.: Does a prototype scaffold/ matrix-attached region (SAR sequence) affect intrinsic nuclear matrix endonuclease specificity? J. Plant Physiol. **141**, 668-672 (1993)
Wirth, M. Schumacher, L. and Hauser, H.: Use of dicistronic transcription units for the correlated expression of two genes in mammalian cells. In: Production of biologicals from animal cells in culture. R. E. Spier, J. B. Griffiths, and B. Meignier (eds.), Butterworth-Heinemann, 338-343 (1991)
Wirth, M., and Hauser, H.: Genetic engineering of animal cells. In: Biotechnology, A. Pühler (ed.) VCH, Weinheim, Vol. **2**, 663-744 (1993)
Wirth, M., and Hauser, H.: Genetic engineering of animal cells. In: Genetic engineering of animals, A. Pühler (ed.). VCH, Weinheim 1-82 (1993)
Wirth, M., Grannemann, R., and Hauser, H.: Screening retroviral helper cells for highly efficient virus production using a combined selection procedure. J. Virology, **68**, 566-569 (1994)

Structural and functional analysis of regulatory genes controlling liver-specific proteins (BIOT CT-910260)

COORDINATOR:
(01) V.I. ZANNIS, Univ. of Crete, Iraklion, Crete, GR
PARTICIPANTS:
(02) V. DE SIMONE, Genetic Engineering Center of Napoli, Napoli, I
(03) R. CORTESE, Istituto di Recherche di Biologia Molecolare, Pomezia, I
(04) G. BRESSAN, Institute of Histology & Embryology, Padova, I
(05) J. CHAMBAZ, Biochimie Lab, URA CNRS 1283, CHU St. Antoine, Paris, F
(06) M. YANIV, Pasteur Institute, Paris, F
(07) G. CILIBERTO, Istituto di Ricerche di Biologia Molecolare, Pomezia, I
(08) M.C. WEISS, Genetic Unit, Dept. Mol. Biol., Pasteur Institute, Paris, F
(09) G. MIGNOT, Lab. Cell Culture, Les Ulis, F
(12) G. BROWNLEE, Chem. Pathol., Univ. of Oxford, Oxford, GB
(13) A.M. ROLLIER, Dept. Biol. Genet., Univ. Studi di Milano, Milano, I

BACKGROUND INFORMATION

The overall objective of this Bridge Project was to study the regulatory proteins and genes which control the synthesis of liver-specific proteins. The liver became the focus of this project because it is the site of synthesis of many proteins that are crucial for normal body functions as well as in diverse pathological states which can be controlled.

OBJECTIVES AND PRIMARY APPROACHES

The specific objectives of the proposed research were:

(1) To isolate and characterize several genes involved in the transcriptional regulation of liver-specific proteins.

(2) To study the structure and the functions of these regulatory proteins by site directed mutagenesis, *in vitro* transcription, and inactivation of the genes involved by homologous recombination in mice.

(3) To identify and characterize the promoter elements and trans-acting factors involved in the transcriptional control of several important liver-specific genes such as apolipoproteins factor IX and albumin genes.

(4) To study the role of extracellular matrix in advanced tissue culture systems on liver functions.

RESULTS AND DISCUSSION

(01 Vassilis Zannis and John Talianidis

I. Transcriptional Regulation of the Human ApoCIII, ApoA-I and ApoA-IV Gene Complex: The ApoCIII Promoter Elements Act as an Enhancer in the Transcription of the Human ApoA-IV Gene

The apoCIII gene is closely liked to the apoA-I and apoA-IV genes (Fig. 1A). Characterization of the activities which bind to the apoCIII promoter showed that HNF-4 binds to the proximal element B (-87 to -72), and can activate transcrip-

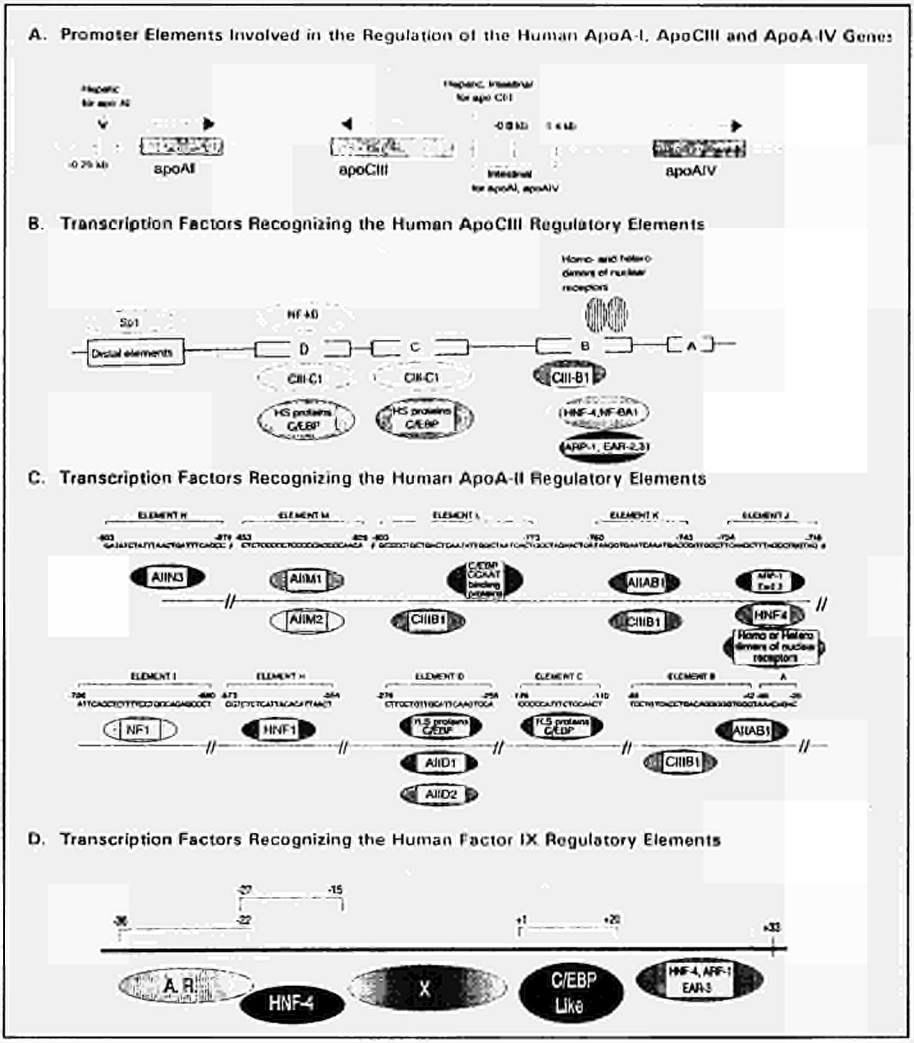


Fig. IA-Dh. Panel A. *Long-distance regulatory elements controlling the expression of human apoCIII, apoA-I, and apoA-IV genes.* ***Panel B.*** *Proximal and distal regulatory elements of the human apoCIII gene and transcription factors which bind to them and regulate transcription.* ***Panel C.*** *Proximal and distal regulatory elements of the human apoA-II gene and transcription factors which bind to them and regulate transcription.* ***Panel D.*** *Proximal regulatory elements of the factor IX promoter and factors which bind to them and regulate transcript*

tion only in the presence of the upstream regulatory elements, F, G, H, I, and J. Element B is also recognized by homo- and heterodimers of nuclear receptors. Elements F, H and I can bind Sp1 or related factors. Element G binds distinct members of the hormone receptor family, like Arp-1 and Ear-3. Two minor activities designated CIII-J1 and CIII-I5, bind to the regulatory elements J and F, and I respectively. The various activities which bind to the apoCIII promoter are

shown in Fig. 1B. Recent experiments indicate that the apoCIII promoter acts as an enhancer and regulates the transcription of the distantly linked apoA-I and apoA-IV genes. *In vitro* footprinting analysis with rat liver nuclear extracts have identified 4 protected regions: A (−22 to +32), B (−42 to −84), C (−120 to 148) and D (−250 to −274). Element A-IV C binds to HNF-4, Arp-1 and Ear-3 with similar affinity (Kd = 4.3-7.7 nM). HNF-4 is the major positive factor responsible for the apoA-IV promoter activity as evidenced by *in vitro* transcription analysis. Although the apoA-IV promoter (−700 to +10) is very active *in vitro*, a very low activity was detected in transfection assays in HepG2 and $CaCo_2$ cells. This low activity was dramatically increased when the apoCIII promoter region was linked to it in tandem or reverse orientation. A series of different constructs containing different combinations of apoCIII promoter elements linked to the −700 to −10 apoA-IV promoter have been generated. Transfection experiments in HepG-2 and CaCo-2 cells indicated that the enhancer activity is localized within the regulatory elements F-J of apoCIII. This region combined with the proximal −700 to −10 apoA-IV promoter region allowed an HNF-4 dependent activation of a reporter gene. Arp-1 and Ear-3 repressed this activation by competing with HNF-4 for the common binding site (A-IVC). These findings suggest that a common set of transcription factors binding to the distal apoCIII regulatory elements F-J is responsible for the tissue specific expression of the apoCIII and apoA-IV geion.

II. Identification of an Indirect Autoregulatory Mechanism Involved in HNF-1/LFB1 Gene Regulation.

HNF-4 is an essential positive regulator of HNF-1. In transient transfection assays, HNF-1 strongly downregulated transcription driven by its own promoter in HepG2 cells. In addition, HNF-1 also repressed the activity of HNF-4 dependent apoCIII and apoAI promoters. The same effect was observed in experiments employing vHNF-1/LFB3. Both HNF-1 and vHNF-1 downregulated HNF-4 activated transcription from intact and chimeric promoter constructs carrying various HNF-4 binding sites implying that they act by impeding HNF-4 binding or activity. DNA binding and cell free transcription experiments however failed to demonstrate any direct or indirect interaction of HNF-1 and vHNF-1 with the above regulatory regions. Both factors repressed HNF-4 induced transcription of the apoCIII and HNF-1 genes in HeLa cells, arguing against the requirement of a hepatocyte specific function. These findings define an indirect negative autoregulatory mechanisms involved in HNF-1 gene expression, which in turn may affect HNF-4 dependent transcription of other liver specific genes such as apoCIII and apoA-I.

(02) Vincenzo DeSimone

I. LFB1/HNF-1 and LFB3/vHNF-1: Sites of Synthesis and Their Role in Liver Differentiation.

During the development of the kidney LFB3/vHNF-1 transcription starts with the induction of the presumptive mesenchymal cells, while LFB1/HNF-1 appears only when the proximal and the distal tubules are detectable. The induction pattern of LFB3/vHNF-1 can be reproduced in the transfilter organ culture system, in which the presumptive kidney mesenchymal tissue can be grown *in vitro* and induced to differentiate by co-cultivation with spinal cord explants. In this system, LFB3/vHNF-1 expression 'bursts' between 24 and 48 hours after induction, when the first morphologic differentiation of the nephrogenic vesicles occur.

II. Analysis of the Regulatory Elements and Factors Controlling the Transcription of LFB3/vHNF-1.

Fusions of the LFB3/vHNF-1 5' proximal regulatory region to a reporter gene combined with deletions, site-specific mutagenesis and DNA binding assays identified the following cis-acting elements and the factors required for transcription of LFB3/vHNF-1 gene.

1) The OCT protein binds adjacent to the TATA box.
2) The liver-specific HNF-4/LFA1 factor binds in the -260 nt region.
3) Homeo- or heterodimers of retinoic acid receptors bind at a site partially overlapping that of HNF-4/LFA1.

III. Functional Analysis of LFB3/vHNF-1.

Mapping of the transcriptional activation domain of the LFB3/vHNF-1 by deletion and nucleotide substitution mutagenesis of the protein coding region showed that the protein region spanning amino acids 350-450 is crucial for transactivation of target genes.

(03) Ricardo Cortese

I. A POU-A Related Region Dictates DNA Binding Specificity of LFB1/HNF-1 by Orienting the Two Extra-Long Homeodomains in the Dimer

The transcription factor LFB1/HNF-1 regulates the hepatocyte-specific transcription of several genes by binding as a dimer to cis-acting elements that match the inverted palindrome GTTAATNATTAAC. The DNA binding domain of LFB1/HNF-1 is characterized by a unique tripartite structure that includes an unusually long homeodomain (domain C), a region related to the POU-specific A-box (domain B) and a short N-terminal dimerization domain (domain A). A recombinant peptide corresponding to the isolated homeodomain of LFB1/HNF-1 binds as a monomer to a half-palindrome binding site, but shows diminished sequence specificity. Domain B, in addition to the homeodomain, is required and is sufficient for proper recognition of LFB1/HNF-1-responsive sites. A protein consisting of only these latter two domains is a monomer in solution, but forms dimers upon binding to DNA. The protein-protein contacts established within the bound dimer restrain the orientation of the two homeodomains with respect to one another, thus contributing in a critical fashion to the recognition of the dyad symmetry-related LFB1/HNF-1 sites. The DNA-dependent dimerization domain (domain A) is required to increase the affinity of DNA binding, but does not influence the dimer geometry.

II. Characterization of the Transactivation Domains of LFB1/HNF-1

Previous *in vitro* experiments suggested the presence of two different regions in the carboxy-terminal portion of the protein responsible for most of the transcription activation potential of LFB1/HNF-1. The first domain (ADI) is located between amino acids 546 and 628 and the second (ADII) between amino acids 281 and 318. Several LFB1/HNF-1 deletion mutants were tested for their ability to induce transcription from LFB1/HNF-1-dependent synthetic promoters in cells of hepatic and nonhepatic origin *in vivo*. The results confirm the importance of ADI and indicate that no activating property can be assigned to ADII *in vivo*. Moreover, a novel glutamine/proline-rich activation domain (ADIII) has been identified between amino acids 440 and 506. The findings indicate that LFB1/HNF-1 transcription activating potential relies on a complex structure and also provide important clues

towards the understanding of the different functions exerted by transcription factors of this family.

III. Secondary Structure and Conformational Stability of the Dimerization Domain of the Transcription Factor LFB1/HNF-1

The structure of a 32 residue dimerization domain of LFB1/HNF-1 in solution has been determined by nuclear magnetic resonance. 3 structurally distinct regions can be distinguished. The N-terminal region from residues 1-6 is extended. Two helical regions span from residues 7-18 and from 23-32. The absence of dipolar effects involving residues more than four positions apart in the sequence excludes the possibilities both of a four-helix bundle formed by two hairpins and of an antiparallel dimer; the domain must therefore be arranged as a parallel dimer formation by kinked monomers. This structural solution presents important differences from the leucine zipper-type structure observed in other transcriptional activators. The thermodynamic parameters associated with the unfolding of the 32 residue long dimerization domain were also determined. The results indicate that the dimerization domain of LFB1/HNF-1 can fold and dimerize independently of the rest of the protein, with a thermodynamic stability comparable to that of a small globular protein.

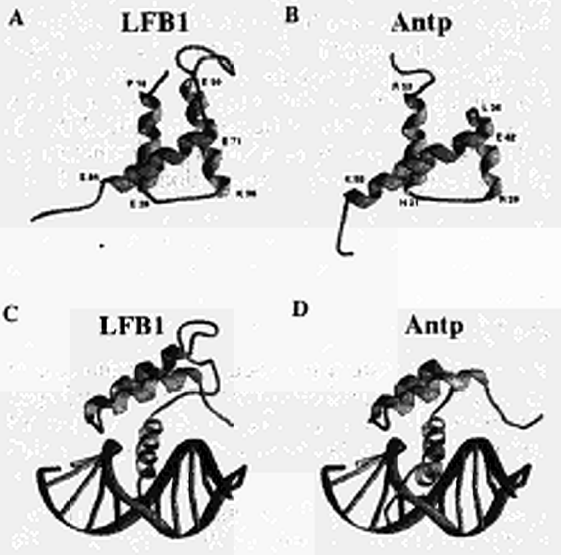


Figure 2. Ribbon plot of the DNA binding domain of LFB1/HNF-1 and the Drosophila transcription factor Antennapedia (Antp) as determined by NMR spectroscopy in solution. Note that the three-dimensional arrangement of the helices is similar in LFB1/HNF-1 (Panel A) and Antennapedia (Panel B). Panels C and D show potential interactions of these protein domains with their cognate sites on the DNA.

IV. The Three-Dimensional Structure of the DNA Binding Domain of LFB1/HNF-1 as Determined by X-Ray Crystallo-graphy & NMR Spectroscopy

The three-dimensional structure of the DNA binding domain of LFB1/HNF-1 has been determined using two different approaches.

a) *Crystallization of a 99 residue protein containing the homeodomain portion of LFB1/HNF-1, and analysis of x-ray diffraction data to 2.8Å resolution.* The x-ray crystallography showed that the topology and orientation of the helices is essentially the same as that found in the engrailed, MAT α2 and Antennapedia homeodomains, even though the LFB1/HNF-1 homeodomain contains 21 more residues. The 21 residue insertion is found in an extension of helix 2 and conse-

quent lengthening of the connecting loop between helix 2 and helix 3. Comparison with the engrailed homeodomain-DNA complex indicates that the mode of interaction with DNA is similar in both proteins, with a number of conserved contacts in the major groove. The extra 21 residues of the LFB1/HNF-1 homeodomain are not involved in DNA binding. Binding of the LFB1/HNF-1 dimer to a B-DNA palindromic consensus sequence requires either a conformational change of the DNA (presumably bending), or a rearrangement of the subunits relative to the DNA.

b) *Nuclear magnetic resonance (NMR) spectroscopy in solution of a polypeptide fragment containing the homeodomain sequence of LFB1/HNF-1.* A recombinant 99 amino acid polypeptide containing this sequence was uniformly labelled with 15N and also doubly labelled with 15N & 13C and analyzed by NMR spectroscopy in solution. This analysis showed that this sequence is approximately 22% homologous to the well known homeodomains, but consists of 81 amino acid residues as compared with 60 residues in 'typical' homeodomains. The NMR structure of this polypeptide contains 3 α-helices comprising residues 18-29, 36-50 and 71-84, a loop formed by residues 30-35, and a long stretch of non-regular secondary structure linking the second and third helices (Fig. 2A). The relative location and orientation of the helices is very similar to that in the Antennapedia (Antp) homeodomain structure, despite the fact that helix 2 is elongated by about one turn (Fig. 2B). The helix-turn-helix motif, which has been shown to comprise the DNA recognition helix in the Antp homeodomain, can readily be recognized in the LFB1/HNF-1 homeodomain, in spite of an extensive modification of the primary structure. The 2 residues of the tight turn in the Antp homeodomain are replaced by a 23 residue linker region between the 2 helices in LFB1/HNF-1, which bulges out from the rest of the molecule and thus enables the formation of a non-classical helix-turn-helix motif. The interactions of these domains with the DNA is shown in Fig. 2C&D.

(04) Giorgio Bressan

Type VI Collagen Function and the Liver

The extracellular matrix is part of a signaling system involved in the maintenance of tissue structure and its remodelling in pathological conditions. Type VI collagen is a pericellular protein with wide tissue distribution. The function of type VI collagen in different tissues including liver has been addressed using various approaches as follows.

a) Introduction of dominant negative mutations in the mouse by microinjection into oocytes of minigenes of the α1(VI) chain carrying small deletions of the collagenous portion. Up to now a total of 11 founders for 2 such mutations have been produced and are being characterized.

b) Expression of type IV collagen during mouse development by RNA blotting. This analysis has shown that α1(VI) mRNA is first detectable in whole embryo extracts at day 11.5. By immunohistochemistry and *in situ* hybridization the gene products were detected in mesenchymal cells and cells derived from somites and lateral plate mesoderm. Their appearance usually coincides with and persists during the late phase of organogenesis and then subsides when the final microscopic structure of the tissue has been reached. In the liver collagen α1(VI) gene products were observed in the perinatal period, when hemopoiesis declines and the organ acquires its final structure with a peak at 12 days after birth. The protein is mainly localized in the perisinusoidal space. In the adult animal the VI collagen

mRNA was not appreciable in the liver. These observations suggest a role for collagen VI in liver morphogenesis.

c) Expression of type VI collagen in cultures of mesodermal cells (fibroblasts, smooth muscle cells & myoblasts). Proliferating cells produced very low levels of mRNA for α1(VI) collagen. However when the cells stopped dividing and differentiated the amount of mRNA increased considerably, peaked after one day and then declined. The protein secreted accumulated at the surface of the cells as fine, dotty aggregates. Evidence of the dependence of collagen VI expression from the cell cycle has also been obtained *in vivo* by comparing the distribution of the protein with that of proliferating cells. The data indicate that type VI collagen may have a role in late organogenesis and tissue remodeling by providing an adhesive matrix which favors the differentiated functions of the cells.

(05) Jean Chambaz and Phillipe Cardot

I. Characterization of Regulatory Elements and Nuclear Factors Involved in the Regulation of the Transcription of the Human ApoA-II Gene

The human apoA-II promoter contains a set of 14 regulatory elements (A to N). Deletion or nucleotide substitution analysis in elements N, L, K and AB affected greatly the hepatic and intestinal transcription. Elements AB, K, and L bind with different affinities to a newly characterized heat-stable factor, CIIIB1, which is a transcription activator of the human apoCIII gene. In addition, elements AB and K bind a heat-labile activity, designated AIIAB1, and element L binds to several CCAAT box binding activities. The CIIIB1 and AIIAB1 activities are very important for transcription and bind contiguously on the AB element, as indicated by methylation interference assays. DNA binding and competition assays and protein fractionation has also shown that the regulatory elements M, D and F bind new activities which have not been identified in apolipoprotein or other liver specific promoters. These activities have been designated AIIM1 and AIIM2 for element M, AIID1 and AIID2 for element D and AIIF2 for element F. In addition, element I binds to activities related to NF1 and elements L, C, D, G, AB and F bind with different affinities to C/EBPα as well as other heat stable activities. Element J is recognized by members of the nuclear receptor family HNF-4, EAR-2, EAR-3 and ARP-1, and element H binds HNF-1. This analysis establishes the organization of several nuclear activities on the human apoA-II promoter (Fig. 1C). The findings suggest that the expression of a liver specific gene such as apoA-II is controlled by a combination of factors and requires synergism of factors bound to proximal and distal regulatory elements. The relative importance of individual factors is determined by the promoter context. Element D exerts negative regulatory control of transcription in HepG2 and $CaCo_2$ cells and binds four activities, designated AIID1, AIID2, AIID3/C/EBP and AIID4. Activities AIID1, AIID2, and AIID4 were purified and characterized further. SDS-PAGE analysis as well as photoaffinity cross-linking of the affinity purified AIID2 showed that it consists of three proteins with molecular mass ranging between 54 and 63 KDa. Amino acid sequence of tryptic peptides obtained from AIID2 protein bands revealed that it is homologous to GABP, an Ets-related protein. Similar analysis showed that affinity purified AIID4 has an apparent molecular mass of 130 KDa. Methylation interference of G residues and permaganate modification of T residues indicated that the binding site of AIID2 and AIID4 were contiguous on element D, whereas the binding site of AIID1 overlaps with the binding sites of both AIID2 and AIID4. Transcription from a minimal promoter containing elements AB, C and D of apoA-II increased 1.5- to 1.6-fold when element D was deleted, as well as by

promoter mutations which eliminated the binding of both AIID1 and/or AIID4 to element D, but permitted the binding of AIID2/GABP. The findings suggest that element D has a negative regulatory role on apoA-II gene transcription when it is occupied by protein AIID1 and/or AIID4. This negative effect is reversed when element D is occupied only by the regulatory factor AIID2/GABP.

II. The Regulatory Region -911 to +29 of the ApoA-II Gene Directs Liver Specific Transcription in Transgenic Mice

A construct containing the -911 to +29 apoA-II promoter region in front of the CAT gene (which directed transcription in HepG2 cells) was utilized to generate transgenic mice. mRNA analysis showed that the -911 to +29 apoA-II region is sufficient to direct transcription in a tissue specific manner, demonstrating expression of the gene only in the liver.

(06) Moshe Yaniv

I. Functional Analysis of HNF-1/LFB1 and vHNF-1/LFB3. Homeodomain Proteins Which Are Involved in Liver Specific Gene Expression

In a series of experiments the regions (residues) of HNF-1/LFB1 protein involved in transcriptional activation has been mapped by deletion analysis and expression studies. Deletion of C-terminal residues resulted in a gradual decrease in the normalized transcriptional activity. Approximately 150 C-terminal residues are important for transcriptional activation. The activity was totally lost by deletion of 236 C-terminal amino acids. The C-terminal deletions increase the stability of the protein. Deletion of approximately 200 amino acids exposed a domain that retains the protein in the cytoplasm. Further deletion that removes the 330 C-terminal residues restores nuclear transport (the DNA binding domain contains a clear nuclear localization signal). The findings indicate that the cytoplasmic retention is probably an active process.

II. Isoforms of HNF-1/LFB1 and vHNF-1/LFB3 Generated by Alternate RNA Processing and Potential Functions

Another aspect concerning the transactivation domains of HNF-1 and vHNF-1 was revealed by a study of alternatively spliced forms of these proteins derived from human liver cDNA library. Three classes of HNF-1 cDNA clones differing in their 3' portions were isolated and have been named HNF-1A (the original form), HNF-1B and HNF-1C. In the last two forms, alternative RNA processing deletes C-terminal sequences of the protein and replaces them with short sequences translated from intronic sequences in the A form. These two forms are more potent activators than the initially isolated HNF-1A clone. RT-PCR studies have shown that HNF-1C is abundant in human fetal kidney and persist to some degree in adult kidney.

Alternative RNA processing generates three forms of vHNF-1 transcripts named vHNF-1A (the original form), vHNF-1B and vHNF-1C. In rat and human an alternative spliced exon is located in the DNA binding N-terminal half of the protein. vHNF-1A which contains this additional exon is a more potent activator than HNF-1B which lacks this exon. Alternative polyadenylation in the human gene generates the third form, vHNF-1C. This form lacks most of the transactivation domain and behaves as a transdominant repressor when cotransfected with HNF-1A or when introduced into human hepatocytes. vHNF-1A and B mRNA is found in liver, kidney, intestine, lung, thymus and ovary with the highest concentration found in liver and kidney. vHNF-1C mRNA is found predominantly in kidney, in fetal intestine and in colon carcinoma cell lines.

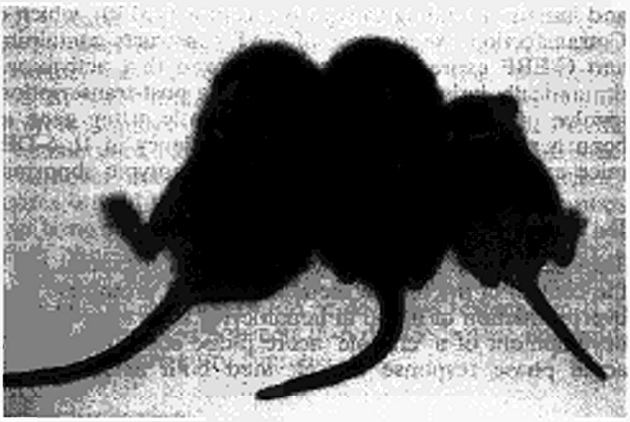


Fig. 3. Inactivation of the HNF-1 gene in mice. The picture shows (left to right) normal control, heterozygous and homozygous deficient mouse for the transcription factor HNF-1A.

III. Inactivation of the HNF-1/LFB1 Gene in Mice

Further insight on the *in vivo* functions role of HNF-1 was obtained by inactivation of the HNF-1 genes of the mouse by homologous recombination. The first exon of HNF-1 was replaced with a β-gal tk neo cassette in ES cells. This mutation was transferred successfully to the germ line. In heterozygotes, HNF-1 expression can be easily followed by X-gal staining and confirmed that the gene is expressed in liver hepatic cells, in the proximal tubules of the kidney, in the base of the intestinal crypts and in some cells of the stomach. In an outbred background animals homozygous for the mutation are born in the expected Mendelian ratio. They are however dwarf, become sick several days after birth, and die soon after weaning (Fig. 3). They suffer from hepatomegaly and exhibit several histological and biochemical defects. However, the precise cause(s) for their morbidity has not been determined as yet. Surprisingly, several of the HNF-1 target genes like albumin, α-antitrypsin or fibrinogen are still expressed. One possible explanation is the upregulation of vHNF-1, the other member of this family in the liver of HNF-1 deficient mice.

(07) Gennaro Ciliberto and Valeria Poli

I. Study of Structure-Function Relationships in IL-6, Aiming at the Future Development of IL-6 Antagonists. Signal Transduction by IL-6 in HepG2 Cells Via C/EBP Family Members

A series of single amino acid variants of IL-6 in residues 176-181 have been generated by *in vitro* mutagenesis based on a three-dimensional model of the protein that was constructed based on the x-ray structure of related cytokines. Analysis of their biological activity in two systems (murine hybridoma IL-6 dependent cells, and human hepatoma cells) and of receptor binding properties of a subset of them indicates that the entire region is involved in forming the receptor binding surface. Remarkably, a mutant was also found with receptor affinity and biological activity 3- to 5-fold higher than wild-type: this molecule is of practical interest, given the potential of using IL-6 in the treatment of neoplastic diseases and thrombocytopenia. Taking into account the available knowledge on the binding of helical cytokines to dimeric receptors the ongoing research is directed towards the generation of IL-6 variants with intact or improved binding to type A receptor

and impaired binding to type B receptor (gp130), which can be used as antagonist. Cotransfection experiments of CAT constructs containing IL-6 response elements and C/EBP expression plasmids showed that activation by IL-6-DBP/C/EBPβ is dramatically induced by IL-6 through a post-transcriptional mechanism which may involve protein phosphorylation. Recently, using gene targeting techniques, mice been generated with homozygous deficiency in IL-6-DBP/C/EBPβ. The deficient mice are viable and do not present phenotypic abnormalities.

II. Pathologies Caused by IL-6 Overproduction

Transgenic mice have been generated which express the human IL-6 cDNA under the control of the mouse metallothionein I promoter. These mice show constitutive overproduction of hIL-6 in hepatocytes soon after birth. This is responsible for the development of a 'chronic' acute phase condition: all the classical markers of the acute phase response are elevated both as mRNA in the liver tissue and a

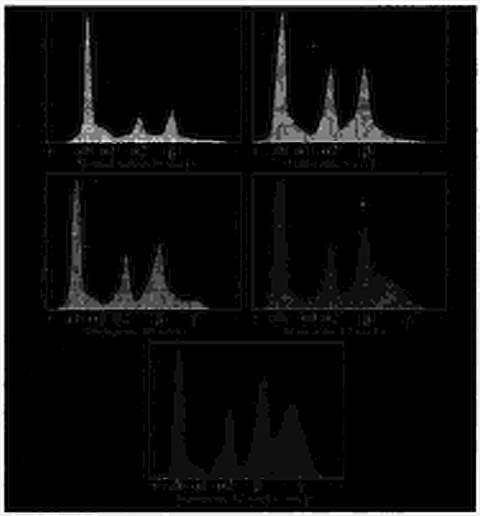


Figure 4. Serum protein profiles of IL-6 transgenic mice expressing the human IL-6 at different ages.

proteins in the serum (Fig. 4). At later ages, IL-6 overproduction leads also to B-cell stimulation and to the development of IgG plasmacytosis and hypergammaglobulinemia. The long-term pathology that develops in these mice at 3-5 months of age has been thoroughly characterized and closely resembles kidney myeloma, the most frequent complication of multiple myeloma in man. Apparently, this pathology is caused by protein overload of the kidney due to the increased serum levels of acute phase reactants and to the hypergammaglobulinemia. Preliminary analysis of the changes in C/EBP expression in the liver of MT-I/IL-6 transgenic mice show substantial changes in the amount of both mRNA and proteins for the various C/EBP proteins. This analysis when completed may establish a precise correlation between the pattern of C/EBP protein expression and the expression of target genes.

III. Inactivation of the IL-6 Gene by Homologous Recombination

Homozygous IL-6 deficient mice, generated by gene targeting, are viable and do not present obvious phenotypic abnormalities, however they display defective

inflammatory response. It is possible that in the absence of this cytokine, some of its functions are taken over by other functionally equivalent molecules. A possible mechanism for such redundancy can be identified by the fact that its effector molecule, gp130, is shared by structurally related cytokines LIF, OM, IL-11. Despite this redundancy however, analysis of bone metabolism revealed a specific phenotype. IL-6 deficient female mice have a normal amount of trabecular bone, but higher rates of bone turnover than control littermates. Estrogen deficiency induced by ovariectomy causes in wild-type animals a significant loss of bone mass together with an increase in bone turnover rates. Strikingly, ovariectomy does not induce any change in either bone mass or bone remodeling rates in the IL-6 deficient mice, indicating that IL-6 plays an important role in the local regulation of bone turnover. Preliminary characterization of the acute phase response in IL-6 deficient mice shows significant differences between IL-6 deficient mice and wild-type controls, depending on the inflammatory stimulus.

(08) Mary C. Weiss

I. Establishing the Transcription Rate of Liver Specific Genes: Cell Phenotype, Promoter Structure and Binding Affinity Modulate Transactivation by HNF-1/LFB1 and LAP

Co-transfection of expression vectors encoding transcription factors with reporter CAT plasmids showed that well-differentiated hepatoma cells produce abundant HNF-1/LFB1 and show high levels of expression of a reporter CAT gene whose expression is directed by the albumin promoter. Addition of exogenous HNF-1 does not enhance the expression of the albumin promoter-reporter gene constructs. In dedifferentiated variants which do not produce endogenous HNF-1, nor do they express the albumin-CAT gene, a strong transactivation of the CAT gene by exogenous HNF-1 is observed. When the albumin HNF-1 binding site is substituted by a sequence of weaker binding affinity, exogenous HNF-1 possesses transactivation potential in all cell lines, whether or not endogenous HNF-1 is abundant. It can be concluded that the threshold of activity of a given transcription factor on a particular promoter can vary depending upon the affinity of the factor for the cognate binding site.

II. HNF-4 and HNF-1 As Well As a Panel of Hepatic Functions Are Extinguished and Reexpressed in Parallel in Chromosomally Reduced Rat Hepatoma-Human Fibroblast Hybrids

Rat hepatoma-human fibroblast hybrids containing 8-11 human chromosomes show pleiotropic extinction of thirteen out of fifteen hepatic functions examined. Reexpression of the entire group of functions most often occurs in a single step, and except for one discordant subclone, correlates with loss of human chromosome 2 as monitored by chromosome 'painting' with immunofluorescence (Fig. 5). The extinguished cells and their reexpressing derivatives have been examined for the expression of seven liver-enriched transcription factors. C/EBP, LAP, DBP, HNF-3 and vHNF-1/LFB3 expression are not systematically extinguished in parallel with the hepatic functions. However, HNF-1/LFB1 and HNF-4 show a perfect correlation with phenotype: these factors are expressed only in the cells showing pleiotropic reexpression. Since recent evidence indicates that HNF-4 controls HNF-1 expression, it can be proposed that the HNF-4 gene is the primary target of the pleiotropic extinguisher.

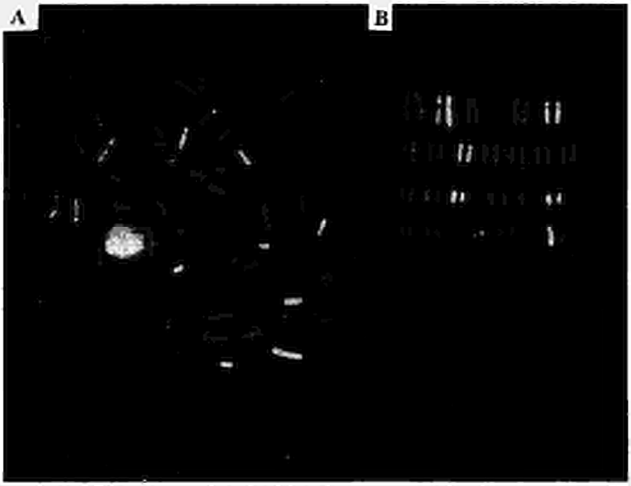


Figure 5. Chromosome painting of a diploid human male metaphase using fluorescent in situ hybridization. Panel A (left), going clockwise from the interphase nucleus, the 'painted' chromosomes (appear-ing white) can be identified: at 9 o'clock, a group comprising 5, 15 and 21 (the latter two are nearly touching); at 11 o'clock, a 2; a 1-2 o'clock, 5, 21 and X; at 3 o'clock, moving outward 18, 18 and 8; at 5-6 o'clock, 8, 2 and 15. Panel B shows the yellow painted human chromosomes 2, 5, 8, 15, 18, 21 and X. Note that both copies of 21 are associated with other acrocentric chromosomes. The numbering of chromosomes in Panel B goes from left to right in rows 1-4.

III. Expression Patterns of vHNF-1 and HNF-1 in Early Post-Implantation Embryos Suggest Distinct and Sequential Developmental Roles

During development, the putative target genes of HNF-1 and vHNF-1 are initially expressed in the visceral endoderm of the yolk sac and subsequently in fetal liver. The expression pattern of both factors was analyzed both *in vitro* during differentiation murine F9 embryonal carcinoma cells and *in vivo* during early post-implantation mouse development. The differentiation of F9 cells into either visceral or parietal endoderm is accompanied by a sharp induction in vHNF-1 mRNA and protein. By contrast, only low levels of aberrantly sized HNF-1 transcripts, but not DNA binding protein, are found in F9 cells and its differentiated derivatives. At 6-7.5 days of gestation high levels of vHNF-1 mRNA are present in the visceral extra-embryonic endoderm, which co-localize with transcripts of the transthyretin gene. HNF-1 transcripts are first detected in the yolk sac roughly two days later, after the developmental onset of transcription of target genes. In addition, two alternative spliced isoforms of vHNF-1 mRNA, vHNF-1A and vHNF-1B, are expressed in embryonic and adult tissues. These data suggest that vHNF-1 participates in the initial transcriptional activation of the target genes in the visceral endoderm of the yolk sac, whereas the later appearance of HNF-1 could be required for maintenance of their expression.

IV. Constancy of Expression of the Protein Kinase A Regulatory Subunit R1αα in Hepatoma Cell Lines of Different Phenotypes

Recently, it was demonstrated that TSE1 (tissue-specific extinguisher) corresponds to R1α, a regulatory subunit of protein kinase A. High expression of R1α characterizes fetal-type BW1J hepatoma cells where the neonatal target genes are silent. This R1α is active in trans to extinguish these genes in hybrids between BW1J and Fao adult-type rat hepatoma cells. Reexpression of the target genes is correlated with loss of R1α and/or overexpression of the mRNA for the transcription factors HNF-4 and HNF-3α. Phenylalanine hydroxylase is negatively regulated by R1α. In BW1J cells where expression of phenylalanine hydroxylase has been activated (i.e. after 5-azacytidine treatment) no downregulation of R1α expression occurs. This suggests that an independent mechanism overcomes R1α repression. Finally, dedif-

ferentiated derivatives of the adult-type rat hepatoma cells express neither the R1α target genes nor the R1α gene itself. Thus, in three different situations where modulation of R1α expression could be anticipated, it fails to occur. It is concluded that somatic cells in culture show cis-heritability of the R1α gene, even under conditions where its effect is selected against.

(09) George Mignot

The task of T.M. Innovation was to optimize the culture condition required for culture of new types of hepatic cell lines, and the upscaling of hepatocyte cultures using macroporous carriers (gelatin or glass spheres) to achieve high cell densities with a combination of high microcarrier concentration and perfusion system.

(10) George Brownlee

Characterization of the Clotting Factor IX Promoter

A total of 21 mutations have been reported in the 1993 database on the promoter of factor IX, defining 13 molecularly unique point mutants, clustered over a discrete region of the promoter from nucleotides −26 to +13. DNA binding and competition assays have identified the binding site for a C/EBP-like protein which recognizes the region +1 to +18 of the promoter. Studies of a patient, haemophilia B Brandenburg (a G→C mutation at −26), and others including the classic promoter mutant at −20 (Haemophilia B Leyden) allowed the identification of an HNF-4 site (−25 to −15) overlapping with an androgen responsive element (ARE) (−36 to −22) (Fig. 1D). Patients with Hemophilia B promoter mutations have a more-or-less severe bleeding disorder, requiring treatment during childhood, but that shortly after puberty the patients improve and no longer need the treatment. Brandenburg patients failed to improve after puberty. The reason for this turned out to be the presence of an androgen responsive element (ARE) in the factor IX promoter, which the mutation in Brandenburg disrupts. Transient transfection assay in HepG2 cells showed that this ARE functions in these cells, requiring the androgen receptor and testosterone for activity. When this ARE is mutated as in Brandenburg it fails to respond. These functional assays have been confirmed by gel shifts with the DNA binding domain of a recombinant androgen receptor. The reason for the reduced expression of factor IX before puberty is that there is a LF-A1/HNF-4 binding site partly overlapping with the ARE (Fig. 1D). DNA binding with purified LFA-1/HNF-4 and competition assays with wild type and mutant oligonucleotides showed that the wild-type factor IX promoter bound LFA-1/HNF-4, whereas both the Brandenburg mutation (at −26) and the classic Leyden mutation (at −20) fail to bind, or to compete efficiently. A new British family, Haemophilia B Liverpool, with mild haemophilia B associated with a −6 G→A mutation in the factor IX promoter has been characterized and carries the same mutation previously observed in Haemophilia B High Wycombe. Very recently, a third British family with this same −6 mutation has been characterized. Preliminary evidence now favors the hypothesis that, despite the fact that this mutation occurs in a CG dinucleotide, a known 'hot spot' for mutation, a founder effect is the probable explanation of the repeat occurrence of this mutation in the 3 British families. Two additional regions, x & y, of the factor IX promoter suspected of being involved in its regulation have been analyzed. The nature of the factors binding to the 'x' region has not been elucidated. However the 'y' region appears to bind HNF-4 and/or members of the ARP-1, EAR-2, COUP-TF family of transcription factors (Fig. 1D).

(11) Anna Maria Rollier

I. Promoter Elements of the Bovine Serum Albumin (BSA) Promoter

Transient expression assay in rat hepatoma H4II showed the BSA promoter has the same tissue specific expression levels as those observed using the highly homologous rat and mouse promoters. A negative acting region was also observed at −230/+170. DNAse I footprinting identified 7 protected sites from −31 to −213 with similar patterns to those seen in rat and mouse promoters. An additional site, observed from −182 to −213, plays a negative regulatory role.

II. Regulation of Albumin Gene Expression in Hepatoma Cells of Fetal Phenotype: Dominant Inhibition of HNF-1/LFB1 Function and Role of Ubiquitous Transcription Factors

Two widely used hepatoma cells lines, mouse BW1J and human HepG2, are characterized by the expression of fetal and adult liver functions, including serum albumin, but express reporter genes driven by the albumin promoter at very low levels compared with highly differentiated hepatoma cells. The low albumin promoter activity in BW1J cells has been studied in order to understand the differences in liver gene regulation between fetal and adult cells. Addition of the albumin upstream enhancer, or any other fragment of the albumin gene, failed to modify expression of the transfected promoter in BW1J cells. Analysis of cis elements of the albumin promoter showed that, in contrast to highly differentiated II4II cells, in BW1J cells the activity depends largely on ubiquitous transcription factors. Both BW1J and HepG2 cells produce the liver-enriched transcription factor HNF-1; dimerization and DNA binding properties are identical to those of liver HNF-1, yet the protein fails to show the anticipated transcriptional stimulatory activity. A transfected HNF-1 expression vector strongly transactivates the albumin promoter in HepG2, but only weakly in BW1J cells, and in hybrids (BW1J x Fao), indicating that the inefficient HNF-1 function is dominant. It is concluded that hepatoma cells of the fetal phenotype are deficient in the use of HNF-1 to drive transcription of the albumin gene, and that they harbor a dominant modulator of HNF-1 function.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

1) Further insight into the regulatory proteins which control the transcription of the apoA-II, apoCIII albumin, factor IX, HNF-1 and vHNF-1 genes, including

a) characterization of the regulatory elements and the factors which control the above genes. The regulatory elements and the factors bound to the factor IX promoter provide molecular explanation of the hemophilia B syndromes which are self-corrected after puberty.

b) Identification of tissue-specific enhancer in the human apoCIII and apoA-II genes.

c) cloning of new members of the HNF-1 family (HNF-1B, HNF-1C) and vHNF-1 (vHNF-1B, vHNF-1C) as well as their tissue distribution and developmental patterns of expression.

d) Evidence of potential interactions between HNF-1 and HNF-4 may lead to negative autoregulation of HNF-1 and other HNF-4 responsive genes.

e) Demonstration of dominant suppression of HNF-1 functions in hepatocytes of fetal phenotype.

f) Potential role of HNF-4 as the target of a pleiotropic extinguisher of liver-specific functions (TSE-1).

g) Evidence that the IL-6 mediated activation of C/EBPβ in HepG2 cells results from post-translational modification possibly involving phosphorylation.

2) Elucidation of the significance of the homeodomain, dimerization domain, and POU-specific A-box for the binding specificity of the liver enriched factor HNF-1/LFB1. b) Localization of domains or residues required for transcriptional activation and possibly cytoplasmic retention of HNF-1 in the C-terminal region.

3) Elucidation of the three-dimensional structure of the DNA binding domain of LFB1/HNF-1 by x-ray crystallography and NMR spectroscopy: This analysis showed that the three-dimensional folding of the DNA binding domain of the mammalian homeodomain protein HNF-1 and its contact points with the DNA are remarkably similar to those of drosophila homeodomain proteins such as Antennapedia. The mammalian homeodomain has an extra 21 residues which form a loop between helices 2 and 3.

4) Derivation of a model for the most probable secondary and tertiary structure of IL-6. Analysis of the functional domains of IL-6 mutants generated by *in vitro* mutagenesis, based on this model, has the potential to generate antagonist of IL-6 with increased affinity for the IL-6 receptor which can be used for treatment of neoplastic diseases and thrombocytopenia.

5) Generation by homologous recombination of mouse lines deficient in HNF-1 and C/EBPβ: These animals provide insights on the *in vivo* functions of these transcription factors in liver and other tissues. Mice with homozygous deficiency for HNF-1 are dwarf, become sick, and die soon after weaning. Mice with homozygous deficiency in C/EBPβ develop normally and do not display obvious phenotypic abnormalities.

6) Generation of animal models either overexpressing or deficient in IL-6 as well as mice with dominant negative mutations in type VI collagen. Mice over-expressing IL-6 develop kidney myeloma. Mice deficient in IL-6 display defective inflammatory response and are protected from bone loss caused by estrogen deficiency. Mice with mutations in collagen (VI) are currently being analyzed.

7) New insights into the ability of the tissue-specific extinguisher 1 (TSE1) which corresponds to the protein kinase A regulatory subunit R1α to repress expression of liver-specific genes: R1α repression can be overcome by other mechanisms.

Wider considerations

The trans-national BRIDGE project entitled 'Structural and Functional Analysis of Regulatory Genes Controlling Liver-Specific Expression' involved 12 laboratories from 4 European countries. The liver was the focus of the present project because it is the site of synthesis of many proteins that are crucial for normal body functions as well as in diverse pathological states. The project involved the study of the regulatory proteins and genes which control the synthesis of some representative liver-specific proteins, such as apolipoproteins, clotting factors, albumin and acute phase proteins that are produced in response to inflammation. The broader goal of this project was to understand what are the molecular mechanisms which permit the genes encoding for these proteins to switch on and off in normal and pathological states. The basic knowledge that has emerged from this project may have several industrial and pharmaceutical applications as follows:

1) Generation of improved hepatocyte cultures which produce large quantities of liver specific proteins of pharmacological importance such as factor IX.

2) Generation of improved hepatocyte cultures which can also be used for toxicology tests, currently performed with livers of animals. This approach will save animal lives and will allow cost-effective and rapid screening of toxic substances.

3) Generation of laboratory animals with partial or total deficiency in regulatory proteins which control the synthesis of liver specific proteins. This approach provides experimental animal models of specific diseases which can be used to explore new modalities of treatment of such diseases. Animals deficient in HNF-1, IL-6 and C/EBPβ have been produced.

4) Design of new types of drugs which are based on the three-dimensional structure of a protein and its interaction with DNA or other proteins or activating ligands. Such information has been obtained for HNF-1/LFB1 and IL-6.

5) Some of the hepatic regulatory proteins may play a role in liver differentiation and regeneration. Drugs which modify the action of these proteins may be used to control liver regeneration.

MAJOR COOPERATIVE LINKS

All the participants of the Bridge Project met three times: once in Sorrento, Italy (June 1-6, 1991) and twice in Heraklion, Crete (September 10-12, 1992 and September 9-11, 1993). The participants had formal presentations of their work, as well as informal discussions on the overall directions of the project. During these meetings they considered practical and industrial implications of the basic knowledge acquired from the study of the regulatory proteins which control liver-specific genes. Dr. Yaniv visited the laboratory of Dr. Zannis once (in 1991). Dr. Zannis visited the laboratory of Dr. Yaniv and Dr. Weiss once (in 1991), the laboratory of Dr. Cortese and Dr. Ciliberto twice (in 1991 and 1992) and the laboratory of Dr. Chambaz twice (in 1991 and 1992). Several of the participants met at the meeting 'Regulation of Liver Gene Expression' in Cold Spring Harbor in 1992. Drs. Zannis, Chambaz and Cardot met at the American Heart Association Meetings in the USA in 1991, 1992 and 1993. Dr. Anna Maria Rollier worked two months in 1993 in the laboratory of Dr. Yaniv and Dr. Weiss.

European dimension

The project brought together several leading European laboratories from four countries with common interest in hepatic gene regulation but with diverse research expertise. Several participating laboratories have made pioneering contributions in the field of eukaryotic gene regulation. The expertise of the laboratories were complementary, thus providing strengths in the fields of molecular biology, cellular biology, and protein chemistry.

The project fostered interactions, not merely among the project leaders, but also among the younger scientists of the groups and thus ensured continuity of interaction among the European scientists. The participation of IRBM in this project both strengthened the overall science and provided a forum for industrial utilization of the research findings.

The collaboration of the European laboratories avoided duplication of efforts, facilitated exchange of scientists, allowed the transfer of materials and expertise, and promoted the integration and upgrading of the European academic institutions.

PUBLICATIONS

Joint Publications

Ramji, D.P., F. Tronche, P. Gallinari, A. Vitelli, G. Ciliberto, and R. Cortese. 1992. Induction of gene transcription by Interleukin-6 (IL-6) requires post-translational modification of IL-6DBP. Serono Symposia Publications from Raven Press (Michel Revel, Ed.), Vol. 88, pp. 52-62 'IL-6: Physiopathology and Clinical Potentials.'

Rollier, A., C.M. DiPersio, S. Cereghini, K. Stevens, F. Tronche, K. Zaret, and M.C. Weiss. 1993. Regulation of albumin gene expression in hepatoma cells of fetal phenotype: Dominant inhibition of HNF-1 function and role of ubiquitous transcription factors. *Mol. Biol. Cell.* **4**: 59-69.

DeMagistris, L., F. Tronche, M.V. Barone, S. Capone, and V. De Simone. 1993. Mapping of the cis-acting signals and retinoic acid responsive elements of the mouse LFB3 promoter. (submitted)

Ramji, D.P., A. Vitelli, F. Tronche, R. Cortese, G. Ciliberto. 1993. The two C/EBP isoforms, IL-6DBP/NF-IL6 and C/EBPδ/NF-IL6β, are induced by IL-6 to promote acute phase gene transcription via different mechanisms. *Nucl. Acid Res.* **21**: 289-294.

Zannis, V.I., P. Cardot, C. Cladaras, E.E. Zanni, J. Chambaz, Y. Talianidis and D. Kardassis. 1994. Transcriptional regulation of the human apolipoprotein genes. In: Transcriptional regulation of liver-specific genes. (F. Tronche and M. Yaniv, eds.) R.G. Landes Publishers, Austin, TX USA (in press)

Cardot, P., D. Pastier, J.M. Lacorte, M. Mangeney, V.I. Zannis and J. Chambaz. 1994. Purification and characterization of nuclear factors binding the negative regulatory element D of human apolipoprotein A-II promoter. *Biochemistry* (in press)

Individual publications

Bach, I., M. Pontoglio, and M. Yaniv. 1992. Structure of the gene encoding hepatocyte nuclear factor 1 (HNF-1) *Nucl. Acid Res.* **20**: 4199-4204.

Bulla, G.A., DeSimone, V., Cortese, R., and Fournier, R.E.K. 1992. Extinction of α1-antitrypsin gene expression in somatic cell hybrids: Evidence for multiple controls. *Genes and Development* **6**: 316-327.

Lazzaro, D., DeSimone, V., DeMagistris, L., Lehtonen, E., and Cortese, R. 1992. LFB1 and LFB3 homeoproteins are sequentially expressed during kidney development. *Development* **114**: 469-479.

Marigo, V., Sitta, A., Volpin, D., and Bressan, G.M. 1992. Mapping of binding sites for monoclonal antibodies to chick tropoelastin by recombinant DNA techniques. *Connective Tissue Res.* **27**: 1-16.

Marigo, V., Daga-Gordini, D., Sitta, A., Volpin, D., and Bressan, G.M. 1992. Effect of monoclonal antibodies to defined regions of tropoelastin on elastogenesis *in vitro*. *Eur. J. Cell Biol.* **57**: 254-264.

Tronche, F., and M. Yaniv. 1992. HNF-1 a homeoprotein member of the hepatic transcription regulatory network. *BioEssays* **9**: 1-9.

Rey-Campos, J., and M. Yaniv. 1992. Regulation of albumin gene expression. Molecular approaches to the study and treatment of human diseases. (T.O. Yoshida, Ed.) Elsevier pp. 239-262.

Angrand, P.O., J.P. Rousset, and M.C. Weiss. 1992. Cell phenotype, binding affinity, and promoter structure modulate transactivation by HNF-1 and LAP. *J. Cell Sci.* **28**: 1083-1092.

Crossley, M., M. Ludwig, K.M. Stowell, P. De Vos, K. Olek, and G.G. Brownlee. 1992. Recovery from Hemophilia B Leyden: An androgen-responsive element in the factor IX promoter. *Science* **257**: 377-379.

Kritis, A.A., E. Ktistaki, D. Barda, V.I. Zannis, and I. Talianidis. 1993. An indirect negative autoregulatory mechanism involved in hepatocyte nuclear factor-1 gene expression. *Nucl. Acid Res.* **21**: 5882-5889.

DeSimone, V. and R. Cortese. 1992. Transcriptional factors and liver-specific genes. *Biochim. et Biophys. Acta* **1132**: 119-126.

Ciliberto, G., V. Colantuoni, R. DeFrancesco, V. DeSimone, P. Monaci, A. Nicosia, D.P. Ramji, C. Toniatti, and R. Cortese. 1993. Transcriptional control of gene expression in hepatic cells. In: 'Gene Expression: General and Cell-Type-Specific.' (M. Karin, ed.) Birkhauser Boston, Chapter 7, pp. 162-242.

Ceska, T.A., M. Lamers, P. Monaci, A. Nicosia, R. Cortese, D. Suck. 1993. The x-ray structure of an atypical homeodomain present in the rat liver transcription factor LFB1/HNF-1 and implications for DNA binding. *EMBO J.* **12**: 1805-1810.

Leiting, B., R. DeFrancesco, L. Tomei, R. Cortese, G. Otting, K. Wuthrich. 1993. The three-dimensional NMR solution structure of the polypeptide fragment 195-286 of the LFB1/HNF-1 transcription factor from rat liver comprises a nonclassical homeodomain. *EMBO J.* **12**: 1797-1803.
Toniatti, C., P. Monaci, A. Nicosia, R. Cortese, G. Ciliberto. 1993. A bipartite activation domain is responsible for the activity of transcription factor HNF-1/LFB1 in cells of hepatic and nonhepatic origin. *DNA Cell Biol.* **12**: 199-208.
Marigo, V., D. Volpin and G.M. Bressan. 1993. Regulation of the human elastin promoter in chick embryo cells. Tissue specific effect of TGF-α. *Biochim. Biophys. Acta* **1172**: 31-36.
Bressan G.M., D. Daga-Gordini, A. Colombatti, I. Castellani and D. Volpin. 1993. Emilin, a component of elastic fibers preferentially located at the elastin-microfibrils interface. *J. Cell Biol.* **121**: 201-212.
Bonaldo P., S. Piccolo, D. Marvulli, D. Volpin, V. Marige and G.M. Bressan. 1993. Murine α1(VI) collagen chain. Complete amino acid sequence and identification of the gene promoter region. *Matrix* **13**: 223-233.
Bach, I. and M. Yaniv. 1993. More potent transcriptional activators or a transdominant inhibitor of the HNF-1 homeoprotein family are generated by alternative RNA processing. *EMBO J.* **12**: 4229-4242.
Sourdive, D., T. Chouard and M. Yaniv. 1993. The HNF-1 C-terminal domain contributes to transcriptional activity and modulates nuclear localization. C.R. Acad. Sci. Paris, Sciences de la Vie/Life Sciences. **316**: 385-394.
Savino, R., A. Lahm, G. Giorgio, A. Cabibbo, A. Tramontano, and G. Ciliberto. 1993. Saturation mutagenesis of the human Interleukin-6 receptor binding site: Implications for its three-dimensional structure. *PNAS* **90**: 4067-4071.
Gregory, B., R. Savino, and G. Ciliberto. 1993. A fast and sensitive colorimetric assay for IL-6 on hepatoma cells based on the production of a secreted form of alkaline phosphatase (SEAP). *J. Immunol. Meth.* **170**: 47-56.
Griffo, G., C. Hamon-Benais, P.O. Angrand, L. West, M. Fox, O. Lecoq, S. Povey, D. Cassio, and M. Weiss. 1993. HNF-4 and HNF-1 are extinguished and reexpressed in parallel with a panel of hepatic functions in chromosomally reduced rat hepatoma-human fibroblast hybrids. *J. Cell Biol.* **121**: 887-898.
Faust, D.M., M. Boshart, T. Imaizumi-Scherrer, G. Schutz, and M.C. Weiss. 1993. Constancy of expression of the protein kinase A regulatory subunit R1a in hepatoma cell lines of different phenotypes. *Cell Growth & Differentiation* **5**: 47-53.
Lecoq, S. Povey, D. Cassio and M. Weiss. 1993. HNF4 and HNF1 as well as a panel of hepatic functions are extinguished and reexpressed in parallel in chromosomally reduced rat hepatoma-human fibroblast hybrids. *J. Cell Biol.* **121**: 887-898.
Shaw, P.M., M. Adesnik, M.C. Weiss, and L. Corcos. 1993. The phenobarbital-induced transcriptional activation of cytochrome P450 genes is blocked by the glucocorticoid progesterone antagonist RU486. *Mol. Pharmacol.* **44**: 775-783
Stowell, K.M, M.S. Figueiredo, G.G. Brownlee, P. Jones and P.H. Bolton-Maggs. 1993. Haemophilia B Liverpool: A new British family with mild haemophilia B associated with a -6 G to A mutation in the factor IX promoter. *Br. J. Haem.* **85**: 188-190.
Laurent-Winter, C., C. Fougere-Deschatrette and M.C. Weiss. 1994. Identification of polypeptides whose presence correlates with retention or loss of an albumin extinguisher chromosome in rat hepatoma-mouse L cell fibroblast microcell hybrids. *Differentiation.* **55**: 225-232.
Shaw, P.M., M.C. Weiss and M. Adesnik. 1994. HNF3 is a major regulator of the CYP2C6 promoter in hepatoma cells. *Molecular Pharmacology* (in press)
Angrand, P.O., C. Coffinier, and M.C. Weiss. 1994. Response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids depends on the integrity of the cAMP pathway. *Cell Growth and Differentiation* (in press).
Marigo V., D. Volpin, G. Vitale and G.M. Bressan. 1994. Identification of a TGF-β responsive element in the human elastin promoter. *Biochem. Biophys. Res. Commun.* **199**: 1049-1056.
Fattori, E., C. Della Rocca, M. Giorgio, B. Dente, L. Pozzi, and G. Ciliberto. 1994. Development of progressive kidney damage and myeloma kidney in IL-6 transgenic mice. *Blood* **83**: 2570-2579.

The development of a genetic and physical map of the porcine genome (Pig Gene Mapping Project — PiGMaP) (BIOT CT-900187)

COORDINATORS:

1. A.L. ARCHIBALD & C.S. HALEY, Roslin Inst. Edinburgh, Roslin, GB

PARTICIPANTS:

2. E.M. TUCKER, Babraham Inst., Cambridge, GB
3. A. VAN de WEGHE, State Univ. of Ghent, Merelbeke, B
4. M. FREDHOLM, Royal Veterin. & Agric. Univ., Copenhagen, DK
   V. NIELSEN, Nat. Inst. of Animal Science, Foulum, DK
5. H. GELDERMANN, Univ. Hohenheim, Stuttgart, D
6. L. OLLIVIER, INRA Station Génét. Quantit. & Appl., Jouy-en-Josas, F
   J. GELLIN, INRA Lab. Génétique Cellulaire, Toulouse, F
   M. VAIMAN, CEA INRA, Lab. Radiobiologie Appliquée, Jouy-en-Josas, F
   P. POPESCU, INRA, Lab. Cytogénétique, Jouy-en-Josas, F
   G. FRELAT, Centre d'Etudes Nucléaires, Fontenay-aux-Roses, F
7. A.A. BOSMA, State Univ. of Utrecht, Utrecht, NL
8. M. GROENEN, Wageningen Agric. Univ., Wageningen, NL
9. W. DAVIES, Norwegian College of Veterinary Medicine, Oslo, N
12. L. ANDERSSON, Swedish Univ. Agricultural Sciences, Uppsala, S
13. I. GUSTAVSSON, Swedish Univ. of Agricultural Sciences, Uppsala, S

ASSOCIATED PARTICIPANTS:

E. SIGNER, University of Leicester, Leicester, GB
R. DAVOLI, Univ. degli Studi di Bologna, Regio Emilia, I
C. MORAN, University of Sydney, Sydney, AUS
M.F. ROTHSCHILD, Iowa State University, Ames, USA
H. HAMEISTER, Univ. Ulm, Ulm D

BACKGROUND INFORMATION

Animal breeders currently assume that economically important traits such as growth, carcass composition and reproductive performance are controlled by an infinite number of genes each of infinitessimal effect. Although this model is known to be unrealistic, it has successfully underpinned the genetic improvement of livestock, including pigs, over recent decades. A map of the pig genome would allow the development of more realistic models of the genetic control of economic traits and the ultimately the identification of the major trait genes. This would allow the development of more efficient marker assisted selection which may be of particular value for traits such as disease resistance and meat quality.

OBJECTIVES AND PRIMARY APPROACHES

The project objectives were as follows:

(1) To produce a genetic map with markers spaced at approximately 20 centiMorgan intervals over at least 90% of the pig genome.

(2) To produce a physical map with at least one distal and one proximal landmark locus mapped on each porcine chromosome arm and also genetically mapped.

(3) To develop a flow karyotype for the pig based on FACS sorted chromosomes.

(4) To develop PCR based techniques to enable rapid genotyping for polymorphic markers.

(5) To evaluate synteny conservation between pigs, man, mice and cattle.

(6) To develop and evaluate the statistical techniques required to analyse data from QTL mapping experiments and to plan and initiate the mapping of QTLs in the pig.

RESULTS AND DISCUSSION

The overall objectives of PiGMaP were to develop a 20 centiMorgan genetic (linkage) map covering 90% of the genome; to produce a physical map with at least one distal and one proximal landmark locus mapped on each chromosome arm and to plan experiments to map the so-called quantitative trait loci (or QTLs).

Genetic (linkage) mapping

The genetic (linkage) map was developed by a group of seventeen laboratories within the collaboration. Reference pedigrees were established in five centres (Scotland (01), France (06), Germany (05), the Netherlands (08) and Sweden (12)). These pedigrees take the form of three-generation families in which grandparents from genetically divergent breeds were crossed to produce the parental (F1) generation which were subsequently intercrossed. In the Scottish, French and Dutch pedigrees the founder grandparental breeds were the Chinese Meishan and the European Large White (Yorkshire). The Swedish and German pedigrees had European Wild Boar and European improved breeds as their grandparents. The karyotypes of some of the founder animals were checked for abnormalities. DNA samples from 118 F2 pigs plus their respective parents and grandparents were distributed to fourteen different laboratories (including one in Australia and one in the United States) for genotyping.

Several classes of molecular genetic markers were developed and used in the genetic mapping studies. First, restriction fragment length polymorphisms (RFLPs) were established by Southern blot analysis. Most of the probes used to reveal RFLPs were expressed sequences in the form of cDNA clones. Both homologous (mainly cDNA) and heterologous (human, rodent, and other mammalian) probes were used to develop these RFLP markers. The diverse origins of the founding breeding stock meant that many loci screened in this manner did indeed prove to be polymorphic. Over fifty such RFLPs have been characterised by the Bologna, Edinburgh, Foulum, Iowa, Oslo, Sydney and Uppsala groups. Second, single strand conformational polymorphisms (SSCPs) have been developed for the 3'-untranslated regions of some genes. These expressed sequences provide the basis for aligning the pig gene map with the maps of other species, in particular with those of the 'map rich' species — man and mouse.

Hypervariable markers based upon both minisatellite and microsatellite loci have also developed. The Merelbeke, Hohenheim and Leicester groups have isolated and characterised sixteen new locus specific minisatellite (VNTR) loci.

Simple tandem repeat (STR) or microsatellite loci are the markers of choice for QTL-mapping studies as they are abundant, evenly distributed and readily genotyped using the polymerase chain reaction (PCR). The structure, number of repeat blocks and chromosomal distribution of porcine microsatellitte loci (dG-dT)n — (dC-dA)n have been found to exhibit the same desirable features as their

human counterparts. Participants 01, 02, 03, 04, 05, 06, 08, 09 and 12 isolated and sequenced several hundred microsatellite loci. Primers for polymerase chain reaction amplification and genotyping have been designed and the highly polymorphic nature of the loci confirmed. Over 150 of these microsatellite loci have been genotyped in the shared mapping pedigrees. Markers with the combined benefits for comparative mapping of expressed sequences and high polymorphic information content (PIC) as provided by microsatellite repeats have been developed by scanning the EMBL and GenBank databases for sequences that have both these attributes. Amongst the loci for which such markers have been developed are *CGA, DAGK, IGF1,* and *SPP1*. Markers based on polymorphic poly(A) tracts adjacent to known genes have also been established. Finally, protein polymorphisms, erythrocyte antigen variants and SLA genotypes have also been determined in some of the reference pedigrees.

Genotypic data on animals in the shared panel of reference families were sent to the Edinburgh laboratory by the typing laboratories and entered into a central relational database built with INGRES software. This database (ResPig) is accessible across the Internet to all the participants who have provided genotyping data. The data have been subjected to two-point and multipoint linkage analysis using the Crimap suite of programs.

A linkage map of the porcine genome has been developed by segregation analysis of 245 genetic markers. Eighty-two of these markers correspond to known genes. Linkage groups have been assigned to all eighteen autosomes plus the X chromosome. As sixty-six of the markers on the linkage map have also been mapped physically, there is significant integration of linkage and physical map data. Six informative markers failed to show linkage to these maps. As in other species, the genetic map of the heterogametic sex (male) was significantly shorter (~16.5 Morgans) than the genetic map of the homogametic sex (female) (~21.5 Morgans). The sex averaged genetic map of the pig was estimated to be ~18 Morgans in length. Mapping information for 59 Type I loci (genes) enhances the contribution of the pig gene map to comparative gene mapping. As the linkage map incorporates both highly polymorphic Type II loci, predominantly microsatellites, and Type I loci it will be useful both for large experiments to map quantitative trait loci and for the subsequent isolation of trait genes following a comparative and candidate gene approach.

Physical (cytogenetic) mapping

Participants in Copenhagen, Toulouse, Uppsala and Utrecht have assigned genes to chromosomes by *in situ* hybridisation techniques.. Although radioactive methods of *in situ* hybridization have been used for some assignments, especially where a short heterologous cDNA probes was used, most experiments now employ fluorescent *in situ* hybridization (FISH). Regional assignments have been made for one hundred and thirty loci, including sitxy-nine anonymous DNA segments. The physical mapping of functional genes to chromosomes is essential to the alignment of the porcine gene map with the maps of other species, in particular with those of humans and mice. The anonymous DNA sequences that have been mapped include cosmid and P1 clones from which microsatellite markers have been developed for linkage studies. The Toulouse laboratory have developed a method termed SLIM-PCR for the isolation of microsatellites from cosmids without subcloning.

Synteny mapping has been effected by analysis of somatic hybrid cell lines. The cell lines available at the outset did not constitute a mapping panel and therefore

further hybrid cell lines were developed by the Copenhagen, Cambridge, Toulouse, Ulm and Utrecht groups. However, analysis of the new pig-rodent somatic hybrid cell lines created for the project, indicate that fragmentation and rearrangement of the porcine chromosomes in the hybrid lines is a problem. The use of alternative fusion partners is being explored. A fluorescent *in situ* hybridization procedure with a porcine SINE (short interspersed elements) probe has been developed which enables the identification of pig chromosomes in pig x rodent hybrid cells. Synteny mapping studies have confirmed the localization of twelve loci. DNA from a hybrid cell line presumed to have retained the p-arm of chromosome 12 as the only pig component has been used to establish a library enriched for sequences from this region of the pig genome.

A physical (cytogenetic) map of the pig genome has been developed by the PiGMaP collaborators. Physical mapping information is available for a total of 142 loci, including 69 anonymous DNA segments. Of these 142 loci, 130 have been regionally mapped on chromosomes by *in situ* hybridization and 12 have been assigned to chromosomes using somatic cell hybrids. Landmark loci have been identified on each chromosome arm of all 18 autosomes, the Y chromosome and the pseudoautosomal region of the X and Y chromosomes.

Flow cytometry

The polydisperse nature of the porcine karyotype allows the chromosomes to be sorted effectively by the use of a dual laser FACS machine. The DNA content of the haploid porcine genome (~2770 Mb) has been estimated from the flow karyotype. The DNA content of individual chromosomes ranging from 295 Mb for chromosome 1 to 68 Mb for chromosome 18 (the Y chromosome is only 47 Mb) has also been estimated.

The groups (02 and 06), that established flow cytometry of pig chromosomes have determined the chromosomal identity of the flow sorted peaks. Material corresponding to a single peak isolated by preparative flow cytometry was amplified by the polymerase chain reaction (PCR) using degenerate primers and fluorescent label incorporated into the PCR product. This fluorescently labelled DNA was then used to probe or paint metaphase chromosomes and thus identify the chromosomal origin of the flow sorted peak. The identity of all twenty peaks, corresponding to the eighteen autosomes plus the X and Y sex chromosomes has been determined by both groups. The resulting fully characterised and confirmed flow karyotype will be useful for future studies directed at particular chromosomes.

Already the flow sorted material has been used to develop chromosome specific libraries for chromosomes 1, 6, 7, 11, 13, 16 and 18. Markers isolated from the chromosome 1, 6 and 13 libraries have been mapped by linkage analysis and *in situ* hybridization. One subchromosomal region specific library for 6p1.1-q1.2 has been established by coincident cloning using DNA isolated from flow sorted chromosome 6 and from a somatic hybrid cell line containing, amongst other porcine chromosome fragments, 6p1.1-q1.2.

Repetitive sequences

Repetitive sequences and the physical components of chromosomes — telomeres and centromeres have also been studied. The characterisation of porcine SINE sequences has proved to be useful in the amplification of the limited amounts of DNA produced by flow cytometry or chromosome scraping and in the characterisation of the chromosome content of somatic hybrid cell lines. A novel genotyping

method named 5' and 3' SINE-PCR was developed to genotype the variable dinucleotide repeats found at the 3' end of about 12% of pig SINE sequences. As the method uses one arbitary primer and a SINE primer no sequencing or cloning is required to develop the markers. Conventional SINE and LINE-PCR products are also polymorphic in the mapping pedigrees. New repeat families have been characterised and mapped to chromosome 3 and 10.

Quantitative trait loci (QTL) mapping

QTL mapping studies have been initiated at each of the five centres that have established mapping populations. The design of effective QTL mapping experiments have been considered. New statistical methods for locating quantitative trait loci have been developed. Already regions of chromosome 4 that influence growth rate, fatness and intestinal length have been identified in a Wild Boar / Large White cross. Fatness, whether measured as the percentage of fat in the abdominal cavity or backfat thickness, appears to be under the control of QTLs that map to the proximal end of chromosome 4. The QTLs for intestinal length and growth rate (from birth to 70 kg) are located distal to the 'fatness' QTLs.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

A genetic (linkage) map of the pig genome covering all 18 autosomes with a resolution of at least 20 centiMorgans has been established. This linkage map which combines both highly polymorphic anonymous DNA markers and polymorphisms within or close to known genes will be an invaluable resource for future studies to map genes influencing economically important traits. A physical (cytogenetic) map with landmark loci identified on each chromosome arm provides a valuable framework for the subsequent isolation of trait genes on the basis of their map positions.

A collaborative study involving the PiGMaP and Nordic pig gene mapping groups established that genes located on pig chromosome 4 have significant effects on growth, fatness and intestinal length. The resolution of pig chromosomes into 20 discrete fractions, corresponding to the 18 autosomes, plus the X and Y sex chromosomes was not only elegant, but also provides the means to focus mapping future mapping studies on particular chromosomes. Already chromosome specific libraries have been developed for such studies.

MAJOR COOPERATIVE LINKS

PiGMaP has been marked by friendly and open collaboration between the participants. The pig gene mapping project (PiGMaP) has brought together key laboratories in seven EC countries and two EFTA countries in a coordinated and collaborative programme to develop a porcine gene map. The initial grouping of sixteen laboratories has been extended to twenty-one in the context of a European Laboratory Without Walls, including laboratories in Australia and the United States. The Genome Mapping Informatic Infrastructure (GEMINI) project initiated in early 1993 under the auspices of the EU Biotech programme has enhanced computing facilities for the PiGMaP collaboration and made the results of PiGMaP more readily accessible to the wider scientific community.

The collaboration between the participating laboratories has operated at several levels. The reference animals necessary for the genetic (or linkage) mapping were provided by five groups. DNA from these animals is shared by a wider grouping

of fourteen laboratories, including laboratories in Australia and the United States. Linkage studies of this shared pool of animals is continuing in a fully collaborative and cooperative manner — data are entered into a central database, linkage analyses performed and the results distributed.

Chromosome specific libraries have been created by combining chromosome sorting in some laboratories with PCR amplification and cloning in other laboratories and the libraries have subsequently been distributed to other participants. Groups isolating markers have collaborated with those using *in situ* hybridization to assign marker loci to chromosomes. For example, cosmids isolated by the Olso group were physically mapped by the Copenhagen laboratory; cosmids and P1 clones characterised by the Edinburgh and Sydney laboratories were mapped by FISH in Toulouse. Markers developed by the linkage mapping laboratories have been sent to Ulm for synteny mapping studies. The development of SINE priming involved a three way collaboration.

Five project meetings have been held in — Edinburgh (April 15th 1991), Toulouse (6th — 7th December, 1991), Copenhagen (27th-29th November, 1992), Ghent (17th-19th June, 1993) and Edinburgh (8th-11th April, 1994). Not only were all participating laboratories represented but also four US groups, one Japanese and one Australian group have attended at least one project meeting with the support of EU funds. Five laboratories took part in a meeting held at Jouy-en-Josas, France in June 1991 to discuss the design of the experiments to map the quantitative trait loci (QTLs). A laboratory coordinators meeting was held in Edinburgh (5th-7th November, 1993) to discuss plans to continue the collaboration with EU support and specifically to map trait genes.

The PiGMaP collaboration has placed Europe in the forefront of pig gene mapping worldwide. The pooling of resources and expertise across Europe means that it has been possible to develop a fully integrated view of the porcine genome.

PUBLICATIONS

Joint publications

Andersson, L., Haley, C.S., Ellegren, H., Knott, S.A., Johansson, M., Andersson, K., Andersson-Eklund, L., Edfors-Lilja, I., Fredholm, M., Hansson, I., Håkansson, J. and Lundström, K. 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**, 1771-1774.

Andersson, L. Archibald, A. L., Gellin, J. and Schook, L. B. 1993. Report of the 1st Pig Gene Mapping workshop (PGM1), 7th August 1992, Interlaken, Switzerland. *Animal Genetics* **24**, 205-216.

Archibald, A. L., Brown, J. F., Haley, C. S., Fredholm, M., Winterø, A. K., Coppieters, W., Van de Weghe, A., Signer, E., Larsen, N. J., Nielsen, V. H., Johansson, M. and Andersson, L. 1992. Linkage mapping in the domestic pig (Sus scrofa). *Animal Genetics* **23**, Suppl. 1:88.

Archibald, A.L., Brown, J. F., Couperwhite, S,, McQueen, H. A., Nicholson, D., Haley, C. S., Coppieters, W., Van de Weghe, A., Stratil, A., Winterø, A-K., Fredhom, M., Larsen, N. J., Nielsen, V. H., Milan, D., Woloszyn, N., Robic, A., Dalens, M., Riquet, J., Gellin, J., Caritez, J. C., Hue, D., Burgaud, G., Ollivier, L., Bidanel, J-P., Vaiman, M., Renard, C., Geldermann, H., Davoli, R., Ruyter, D., Verstege, E. J. M., Groenen, M. A. M., Davies, W., Høyheim, B., Keiserud, A., Andersson, L., Ellegren, H., Johansson, M., Marklund, L., Miller, J. R., Anderson Dear, D. V., Signer, E., Jeffreys, A. J., Moran, C., Le Tissier, P., Muladno, Rothschild, M. F., Tuggle, C. K., Vaske, D., Helm, J,. Liu, H-C., Rahman, A., Yu, T-P., Larson, R. G., Schmitz, C. B. The PiGMaP Consortium linkage map of the pig *(Sus scrofa)*. *Mammalian Genome* (in press).

Archibald, A. L., Haley, C. S., Andersson, L., Bosma, A. A., Davies, W., Fredholm, M., Geldermann, H., Gellin, J., Groenen, M., Gustavsson, I., Ollivier, L., Tucker, E. M., and Van

de Weghe, A., 1991. PiGMaP — a European initiative to map the porcine genome. *Animal Genetics* **22**, Suppl. 1, 82-83.
Archibald, A. L., Tucker, E. M., Van de Weghe, A., Fredholm, M., Nielsen, V., Geldermann, H., Ollivier, L., Gellin, J., Vaiman, M., Popescu, C. P., Bosma, A. A., Groenen, M., Davies, W., Andersson, L., Gustavsson, I., Signer, E., Davoli, R., 1992 The development of a genetic and physical map of the porcine genome (Pig gene mapping project — PiGMaP). Biotechnology Research for Innovation, Development and Growth in Europe (1990-1993) Progress Report 1992. Commission of the European Communities DGXII [ISBN 92-826-3995-5] , pp 220-224.
Bosma, A. A., de Haan, N. A. and Macdonald, A. A., 1991. The current status of cytogenetics of the Suidae: a review. *Bongo Berlin* **18**, 258-272.
Brascamp, E. W., Haley, C. S., Fries, R., Grosclaude, F., Hanset, R. and Womack, J. 1991. Report of Working Group: Genome mapping and association studies. *Proceedings of the 42nd Meeting of the European Association of Animal Production (Berlin)* Vol I, 3 (abstract).
Chardon, P., Peelman, L., Nunes, M., Renard, C. and Renard, M. 1994. Preservation of the local genetic organization within the swine major histocompatibility complex classIII region. *Proceedings of the International Conference on Comparative Gene Mapping in Terrestrial and Aquatic Vertebrates*, Oslo 7-11 February. pp. 51 (abstract).
Chaudhary, R., Chowdhary, B.P., Harbitz, I. and Gustavsson I. 1993. Localization of the ceruloplasmin (CP) gene to the q32-q33 bands of chromosome 13 in pigs by *in situ* hybridization. *Hereditas* **119**, 7-10.
Chowdhary B.P., Johansson M., Gu F., Bräuner-Nielsen P., Tomkinsson B., Andersson L. and Gustavsson I. 1993. Assignment of the linkage group EAM-TRYP2-TPP2 to chromsome 11 in pigs by *in situ* hybridization mapping. *Chromosome Research* **1**, 175-179.
Chowdhary, B. P., Harbitz, I., Davies, W. and Gustavsson, I., 1991. Mapping of genes belonging to the halothane linkage group in pigs using *in situ* hybridization. *Génétique, Sélection Évolution* **23**, Suppl. 1, 96s-99s.
Chowdhary, B. P., Thomsen, P. D., Ellegren, H., Harbitz, I., Andresson, L. and Gustavsson, I. 1992. Precise localization of some genetic markers in pigs using non-radioactive *in situ* hybridzation. *Animal Genetics* **23**, Suppl 1: 92.
Couperwhite, S., Hemmings, B. A. and Archibald, A. L. 1992. A *Bam* HI RFLP at the locus encoding the 65 kDa regulatory subunit of porcine protein phosphatase 2A (PPP2ARB). *Animal Genetics* **23**, 568.
Couperwhite, S., Kato, Y. and Archibald, A. L. 1992. A *Taq*I RFLP at the porcine thyroid stimulating hormone β-subunit locus (TSHB). *Animal Genetics* **23**, 567.
Davies, W., Høyheim, B., Chaput, B., Frelat, G. and Keiserud, A. 1992. Rapid isolation of porcine chromosome 13 specific microsatellites. *Animal Genetics* **23**, Suppl. 1:90-91.
Echard, G., Yerle, M., Milan, D., Gellin, J. and Archibald, A. L. 1993. The gene map of the pig *(Sus scrofa domestica L.)* 2N = 38. *Genetic Maps*, Sixth Edition, Edited by S. J. O'Brien, published by Cold Spring Harbor Laboratory Press. pp4.240-4.249.
Edfors-Lilja, I., Ellegren, H., Winterø, A.K., Ruohonen-Lehto, M., Fredholm, M., Gustafsson, U., Juneja, R.K. and Andersson, L. 1993. A large linkage group on swine chromosome 7 including the MHC class I, class II (DQB) and class III (TNFB) genes. Immunogenetics **38**, 363-366.
Ellegren, H., Chowdhary, B.P., Johansson, M., Marklund, L., Fredholm, M., Gustavsson, I. and Andersson, L. 1994. A primary linkage map of the porcine genome reveals a low rate of recombination in the pig. *Genetics* **137**, 1089-1100.
Ellegren, H., Fredholm, M., Edfors-Lilja, I., Winterø, A.K. and Andersson, L. 1993. Conserved synteny between pig chromosome 8 and human chromosome 4 but rearranged and distorted linkage map. *Genomics* **17**, 599-603.
Ellegren, H., Johansson, M., Chowdhary, B.P., Marklund, S., Ruyter, D., Marklund, L., BrÄuner-Nielsen, P., Edfors-Lilja, I., Gustavsson, I., Juneja, R.K. and Andersson, L. 1993. Assignment of 20 microsatellite markers to the porcine linkage map. *Genomics* **16**, 431-439.
Fredholm, M., Winterø, A.K., Christensen, K., Nielsen, P.B., Davies, W. and Archibald, A.L. 1993. Characterization of twenty-four porcine (dA-dC)n — (dT-dG)n microsatellites: genotyping of unrelated animals from four breeds and linkage studies. *Mammalian Genome* **4**, 187-192.

Fredholm, M., Winterø, A. K., Christensen, K., Thomsen, P. D., Nielsen, P. B., Kristensen, B., Juneja, R. K., Archibald, A. L and Davies, W. 1992. Mapping of genetic markers in a porcine backcross family. *Animal Genetics* **23**, Suppl. 1: 84.
Frengen, E., Davies, W., Kran, S., Thomsen, P. D., Kristensen, T. and Miller, J. R. 1991. Specific amplification of porcine DNA from pig-rodent hybrid cell lines using the polymerase chain reaction and primers from a porcine short interspersed element. *Animal Genetics* **22**, Suppl 1: 88-89.
Frengen, E., Thomsen, P. D., Schmitz, A., Frelat, G., Kran, S. and Davies, W. 1992. Isolation of probes from the p1.1-q1.2 region of porcine chromosome 6 using coincidence cloning. *Animal Genetics* **23**, Suppl. 1:70.
Frengen, E., Thomsen, P., Kristensen, T., Kran, S., Miller, R. and Davies, W., 1991. Porcine SINEs: chracterization and use in species-specific amplication. *Genomics* **10**, 949-956.
Frengen E., Thomsen, P.D., Schmitz, A., Frelat, G. and Davies, W. 1993. Isolation of probes from the p1.2-q1.2 region of pig chromosome 6 using coincidence cloning. In: *Detailed physical mapping in the porcine halothane linkage group*. E. Frengen, PhD Thesis. University of Oslo, Norway, ISBN 82-90934-08-4.
Gu, F., Harbitz, I., Chowdhary, B. P., Bosnes, M. and Gustavsson, I. 1992. Localization of the hormone sensitive lipase (LIPE) and insulin receptor genes in pigs by *in situ* hybridization. *Animal Genetics* **23**, Suppl. 1. 92-93..
Gu, F., Harbitz, I., Chowdhary, B. P., Bosnes, M. and Gustavsson, I. 1992. Chromosomal localization of the hormone sensitive lipase (LIPE) and insulin receptor (INSR) gene in pigs. *Hereditas* **117**, 231-236.
Gu, F., Harbitz, I., Chowdhary, B. P., Davies, W. and Gustavsson, I., 1992. Mapping of the porcine lipoprotein lipase (LPL) gene to the chromosome 14q12-q14 bands by *in situ* hybridization. *Cytogenetics and Cell Genetics* **59**, 63-64.
Haley, C. S. and Archibald, A. L.; Andersson, L., Bosma, A. A., Davies, W., Fredholm, M., Geldermann, H., Gellin, J., Groenen, M., Gustavsson, I., Marmirolli, N., Nielsen, V. H., Ollivier, L., Popescu, P., Signer, E., Tucker, E. M., Vaiman, M., and van de Weghe, A., 1991. The pig gene mapping project (PiGMaP), ELWW leaflet, Commission of the European Communities.
Haley, C. S., Archibald, A. L., Andersson, L., Bosma, A. A., Davies, W., Fredholm, M., Geldermann, H., Groenen, M., Gustavsson, I., Ollivier, L., Tucker, E. M., and Van de Weghe, A., 1990. The Pig Gene Mapping Project — PiGMaP. *4th World Congress on Genetics Applied to Livestock Production.* **XIII**, 67-70.
Haley, C. S., Knott, S. A. and Elsen, J.-M. 1994. Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195-1207.
Harbitz, I., Chowdhary, B.P., Kran, S. and Davies, W. 1993. Characterization of a porcine glucosephosphate isomerase processed pseudogene at chromosome 1q16-q17. *Mammalian Genome* **4**, 589-592.
Hausmann, M., Dolle, J., Arnold, A., Stepanow, B., Wickert, B., Boscher, J., Popescu, C. P. and Cremer, C. 1992. Development of a two parameter slit-scan flow cytometer for screening of normal and aberrant chromosomes: Application to a karyotype of *Sus scrofa domestica. Optical Engineering* **31**, 1463-1469.
Høyheim, B., Keiserud, A. and Thomsen, P. D. 1992. Direct sequencing and *in situ* hybridization of cosmids containing microsatellites. *Animal Genetics* **23**, Suppl. 1:91.
Høyheim, B., Keiserud, A., Chaput, B., Schmitz, A., Frelat, G. and Davies, W. 1992. Rapid isolation of porcine chromosome 6-specific microsatellies. *Animal Genetics* **23**, Suppl. 1:91.
Langford, C.F., Telenius, H., Miller, N.G.A., Thomsen, P.D. and Tucker, E.M. 1993. The preparation of chromosome specific paints and the complete assignment of chromsomes in the pig flow karyotype. *Animal Genetics* **24**, 261-267.
Marklund, L., Winterø, A.K., Thomsen, P.D., Johansson, M., Fredholm, M., Gustavsson, U. and Andersson, L. 1993. A linkage group on pig chromosome 4 comprising the loci for blood group L, GBA, ATP1B1 and three microsatellites. *Animal Genetics* **24**, 333-338.
Marklund, L., Ellegren, H., Johansson, M., Gustavsson, U., Nielsen, P. B., Juneja, R. K., Edfors-Lilja and Andersson, L. 1992. Linkage analysis in the pig using a cross between the European wild pig and a Large White breed. *Animal Genetics* **23**, Suppl 1:89.
McQueen, H. A., Gilmour, P. S., Yerle, M. and Archibald, A. L. 1993. A novel porcine P1 library for integration of the pig map. In: *17th International Congress of Genetics* , Birmingham, 15-21 August 1993, Volume of Asbstracts pp187.

McQueen, H. A., Yerle, M. and Archibald, A. L. 1994. P1 clones for pig genome map integration. In: Eukaryotic Gene Expression — Eighth Tenovus (Scotland) Symposium, Glasgow, 23-25 March, 1994. University of Glasgow, pp 18, (abstract).
Mellink, C. H. M., Lahbib-Mansais, Y., Yerle, M. and Gellin, J. 1993. Localization of four new markers to pig chromosomes 1, 6 and 14 by radioactive *in situ* hybridization. *Cytogenetics and Cell Genetics* **64**, 256-260.
Mellink, C. H. M., Lahbib-Mansais, Y., Yerle, M. and Gellin, J. 1994. Mapping of the regulatory type I alpha and catalytic beta subunits of cAMP-dependent protein kinase and interleukin 1 alpha and 1 beta in the pig. *Mammalian Genome* **5**, 298-302.
Miller, J. R. and Thomsen, P. D., 1991. A centromere-specific tandemly repeated DNA sequence from the domestic pig. Human Gene Mapping 11 (1991). Cytogenetics and Cell Genetics **58**, 2130.
Miller, J. R., Dixon, S. C., Miller, N. G. A., Tucker, E. M., Hindkjaer, J. and Thomsen, P. D. 1992. A chromosome 1-specific DNA library from the domestic pig *(Sus scrofa d.). Cytogenetics and Cell Genetics* **61**, 128-131.
Miller, J. R., Hindkjaer, J. and Thomsen, P. D. 1993. A chromosomal basis for the differential organization of a porcine centromere-specific repeat. *Cytogenetics and Cell Genetics* **62**, 37-41.
Popescu, C. P., Boscher, J., Kerboef, D., Hausmann, M., Dolle, J., Wickert, B. and Cremer, C. 1993. Analysis of an abnormal pig chromosome by cytogenetics, flow cytometry and slit scanning. *Proceedings of the 10th European Colloquium on Cytogenetics of Domestic Animals*, Utrecht, August 1992. pp 178, ISBN 90-6159-017-5.
Ruyter, D., Johansson, M., Groenen, M. and Andersson, L. 1994. A porcine polymorphic microsatellite locus (S0098). *Animal Genetics* **25**, 198.
Signer, E. N., Gu, F., Andersson, L. and Jeffreys, A. J. 1994. A pseudoautosomal minisatellite in the pig. *Mammalian Genome* **5**, 48-51.
The PiGMaP Consortium, 1993. Towards a linkage map of the domestic pig. In: 17th International Congress of *Genetics* , Birmingham, 15-21 August 1993, Volume of Asbstracts pp. 194.
Thomsen, P. D., Qvist, H., Andersson, L., Sjöström, H. and Noren, O. 1993. Assignment of the dipeptidylpeptidase (DPP4) gene to pig chromosome 15. *Mammalian Genome* 4, 604-607.
Thomsen, P. D., Winterø, A. K., Bosma, A. A. and Fredholm, M. 1994. A porcine repeat family located on 3p14 and distal to the rRNA gene cluster on chromosome 10. *Mammalian Genome* **5**, 244-246.
Thomsen, P. D., Bosma, A. A., Kaufman, U. and Harbitz, I., 1991. The porcine PGD gene is preferentially lost from chromosome 6 in pig x rodent somatic cell hybrids. *Hereditas* **115**, 63-67.
Thomsen, P. D., Christensen, K. and Davies, W., 1991. Regional localization of repetitive DNA on the porcine Y chromosome. Human Gene Mapping 11 (1991). *Cytogenetics and Cell Genetics* **58**, 2131.
Thomsen, P. D., Poulsen, P. H., Olsen, J., Harbitz, I. and Miller, J. R., 1991. Comparative gene mapping of the chromosomes in the domestic pig. *Animal Genetics* **22**, Suppl. 1, 80.
Tucker, E. M., Langford, C. F., Bouvet, A., Miller, N. G. A., Telenius, H., Carter, N. P. and Thomsen, P. D. Separation of pig chromosomes by flow cytometry and preparation of chromosome-specific probes. Proceedings of the 10th European Colloquium on Cytogenetics of Domestic Animals (edited by A.A. Bosma), Utrecht. pp250-252, ISBN 90-6159-017-5..
Winterø, A.K., Fredholm, M. and Andersson, L. 1994. Assignment of the gene for porcine insulin-like growth factor 1 (IGF1) to chromosome 5 by linkage mapping. *Animal Genetics* **25**, 37-39.
Yerle, M., Lahbib-Mansais, Y., Thomsen, P.D. and Gellin, J. 1993. Localization of the porcine growth hormone gene to chromosome 12p1.2-p1.5. *Animal Genetics* **24**, 129-131.
Yerle, M., Gellin, J., Le Tissier, P., Mellink, C., Boulliou, A. and Moran, C. 1993. Chromosomal localization of 15 cosmids in the pig by fluorescence *in situ* hybridization. *Proceedings of the 10th European Colloquium on Cytogenetics of Domestic Animals*, Utrecht, August 1992. pp. 266-270, ISBN 90-6159-017-5.
Yerle, M., Goureau, A., Gellin, J., Le Tissier, P. and Moran, C. 1994. Rapid assignment of cosmid clones on pig chromosomes by fluorescent *in situ* hybridization. *Mammalian Genome* **5**, 34-37.
Yerle, M., Lahbib-Mansais, Y., Mellink, C., Goureau, A., Pinton, P., Echard, G., Gellin, J., Zijlstra, C., De Haan, N., Bosma, A. A., Chowdhary, B., Gu, F., Gustavsson, I., Thomsen, P. D., Christensen, K., Rettenberger, G. and Hameister, H. The PiGMaP Consortium

cytogenetic map of the domestic pig (Sus scrofa domestica). Submitted to *Mammalian Genome*.
Zijlstra, C., Yerle, M., Bosma, A. A., de Haan, N.A., Milan, D. and Gellin, J. 1993. Selection of somatic cell hybrids for gene mapping studies in the pig by chromosome painting. *Proceedings ofthe 8th North American Colloquium on Domestic Animal Cytogenetics and Gene Mapping*, Guelph, 13-16 July 1993. pp161-163.
Haley, C. S., Archibald, A., Tucker, E, M. and Miller, J. R., 1991. PiGMaP — a pan-European study to map the pig genome. In: AFRC News, July 1991. Agriculture and Food Research Council, Swindon. pp 4-6
Miller, J. R. and Archibald, A. L. 1993. 5' and 3' SINE-PCR allows genetic mapping of interbreed pig families without cloning and sequencing steps. *Mammalian Genome* **4**, 243-246.
Miller, J. R. and Archibald, A. L. 1993. 5' and 3' SINE-PCR genotyping without cloning and sequencing steps. *17th International Congress of Genetics* , Birmingham, 15-21 August 1993, Volume of Abstracts p. 183.
Nunes, M., Dezeure, F., Gellin, J., Chardon, P. and Vaiman, M. 1992. Isolation of four HSP 70 genes in the pig and localization on chromosome 7 and 14. *Animal Genetics* **23**, Suppl. 1:85.
Nunes, M., Yerle, M., Dezeure, F., Gellin, J., Chardon, P. and Vaiman, M. 1993. Genomic organisation and assignment of three HSP70 genes to the pig major histocompatibility complex and a fourth to chromosome 14. *Mammalian Genome* **4**, 247-251.
Ollivier, L., Gellin, J., Popescu, P. and Vaiman, M., 1992. La cartographie du génome porcin: une aide à la sélection. *J. Rech. Porcine en France* **24**, 1-8.
Yerle, M., Schmitz, A., Milan, D., Chaput, B., Monteagudo, L. V., Vaiman, M., Frelat, G. and Gellin, J. 1992. Definite characterization of porcine flow karyotype by chromosome painting after PARM-PCR. *Animal Genetics* **23**, Suppl. 1:88.

Individual publications

Archibald, A. L. and Couperwhite, S. 1994. Genetic mapping of anchored reference loci in the domestic pig *(Sus scrofa)*. In: *International Conference on Comparative Gene Mapping in Terrestrial and Aquatic Vertebrates*. Oslo, Norway, 7-11 February 1994. pp. 48.
Archibald, A. L. and Haley, C. S. 1993. Mapping the complex genomes of animals and man. *Outlook on Agriculture* **22**, 79-84.
Archibald, A. L. and Haley, C. S. The pig gene mapping project (PiGMaP). *Proceedings of the 10th European Colloquium on Cytogenetics of Domestic Animals* (edited by A.A. Bosma), Utrecht. pp. 1-5, ISBN 90-6159-017-5.
Archibald, A. L. and Haley, C. S., 1992. Progress on the porcine genome. *Human Genome '92*, Nice, France, 14-17 October 1992. (abstract).
Archibald, A. L., 1992. Route planning on the DNA map *Pig Farming* Breeding Supplement. (November) pp21-25.
Archibald, A. L., Brown, J. F., Couperwhite, S. and Haley, C. S. 1991. Reference families and markers for genetic mapping in the domestic pig. 11th International Workshop on Human Gene Mapping (HGM11) *Cytogenetics and Cell Genetics* **58**, 2123.
Archibald, A. L., Brown, J. F., Couperwhite, S., McQueen, H. A. and Haley, C. S., 1992. Analysis of the porcine genome. *BioTechForum Europe* **9**, 574-575.
Archibald, A. L., Couperwhite, S. and Haley, C. S. 1991. Genetic mapping in Meishan/Large White pig families. 1st Mammalian Genetics and Development Workshop, *Genetical Research* **58**, 75.
Archibald, A. L., Couperwhite, S. and Haley, C. S. 1991. Markers for genetic mapping in the domestic pig. *7th Tenovus — Scotland Symposium 'Eukaryotic genes and their expression'*. University of Glasgow, 20-22 March 1991.
Archibald, Alan L. and Haley, Chris S. 1992. The pig gene mapping project (PiGMaP). *Genome News* **No 10**, 26-28.
Benkhalifa, M., Bonneau, M., Popescu, C. P., Boscher, J., Boucher, D. and Malet, P. 1992. A method for cytogenetic analysis of boar spermatozoa using hamster oocytes. *Annales Génétique* **35**, 61-64.
Bigi, D., Davoli, R., Dall'olio, S. and Russo, V. 1994. A *Pvu*II polymorphism at the porcine aminopeptidase N (PEPN) locus. *Animal Genetics* **25**, 65.
Bonneau, M., Boscher, J., Delate, J. J., Legault, C. and Popescu, C. P., 1991. Deux nouvelles translocations réciproques impliquant le chromosome 6 chez le porc domestique. *Ann. Génét.* **2**, 65-69.

Bosma, A. A., de Haan, N. A. and Mellink, C. H. M., 1991. Ribosomal RNA genes located on chromosome 16 of the domestic pig. 11th International Workshop on Human Gene Mapping (HGM11) *Cytogenetics and Cell Genetics* **58**, 2124.
Bouvet, A., Konfortov, B. A., Miller, N. G. A., Miller, J. R. and Tucker, E. M. 1992. Identification of porcine chromosomes in bivariate flow karyotypes using various methods. *Animal Genetics* **23**, Suppl 1: 84.
Brown, J. F. and Archibald, A. L. 1993. A porcine polymorphic microsatellite locus *(S0031). Animal Genetics* **24**, 70.
Brown, J. F., McQueen, H., Couperwhite, S. and Archibald, A. L. (1992) Towards integrated expression, genetic and physical maps for the pig. *Animal Genetics* **23**, Suppl. 1: 89.
Buitkamp, J., Ammer, H. and Geldermann, H., 1991. DNA fingerprinting in domestic animals. *Electrophoresis* **12**, 169-174.
Chaudhary, R., 1992. Physical gene mapping in pigs. Localization of the genes for PGD, ALB, TF, CP and CS using *in situ* hybridization. *Master of Science thesis. Publication No 99*, Uppsala, 1992.
Chen, Y. 1994. Analysis and application of VNTRs in pig. Thesis, Hohenheim University, Germany.
Chowdhary, B. P., 1991. Chromosome mapping of some genes in farm animals. *Dissertation*, Uppsala 1991.
Chowdhary, B. P., Ellegren, H., Johansson, M., Andersson, L. and Gustavsson, I. 1994. *In situ* hybridization mapping of the growth hormone receptor (GHR) gene assigns a linkage group (C9, FSA, GHR and S0105) to chromosome 16 in pigs. *Mammalian Genome* **5**, 160-162.
Chowdhary, B. P., Johansson, M., Chaudhary, R., Ellegren, H., Gu, F., Andersson, L. and Gustavsson, I. 1993. *In situ* hybridization mapping and RFLP analysis of the porcine albumin (ALB) and transferrin (TF) genes. *Animal Genetics* **24**, 85-90.
Coppieters, W. 1994. Het gebruik van DNA-sequentiepolymorfismen als genetische merkers voor het varkensgenoom (The use of DNA-sequence polymorphisms as genetic markers for the pig genome). PhD thesis, University of Ghent.
Coppieters, W., Van de Weghe, A., Peelman, L., Depicker, A., Van Zeveren, A. and Bouquet, Y. 1992. Development of multiplex typing of microsatellites in the pig. *Animal Genetics* **23**, Suppl. 1: 70.
Coppieters, W., Van de Weghe, A., Peelman, L., Depicker, A., Van Zeveren, A. and Bouquet, Y. 1993. Characterization of porcine polymorphic microsatellite loci. *Animal Genetics* **24**, 163-170.
Coppieters, W., van de Weghe, A., Peelman, L. and Depicker, A., 1991. Probes for multiallelic loci from a pig library enriched for variable number of tandem repeats (VNTR). *Arch. Internat. de Physiologie de Biologie et de Biophysique* **99**, 352.
Coppieters, W., van de Weghe, A., Peelman, L., Depicker, A., van Zeveren, A. and Bouquet, Y., 1991. Systematic cloning of dinucleotide repeats and their use as polymorphic markers for the pig genome. *Arch. Internat. de Physiologie de Biologie et de Biophysique* **100**, B36.
Coscelli, M. B., Davoli, R., Farnetti, E. and Russo, V. 1994. BglII polymorphism at the porcine arachidonate 12-lipoxygenase (LOG12) locus. *Animal Genetics* **25**, 57.
Couperwhite, S. and Archibald, A. L. 1993. An EcoRV RFLP at the porcine chromogranin A locus (CHGA). *Animal Genetics* **24**, 331.
Couperwhite, S. and Archibald, A. L. 1993. An XbaI RFLP at the porcine growth hormone receptor locus *(GHR). Animal Genetics* **24**, 331.
Dalens, M., Boulliou, A., Milan, D., Riquet, J., Parrou, J. L. and Gellin, J. 1992. Analysis of porcine (TG)n microsatellite sequences. *Animal Genetics* **23**, Suppl. 1:72.
Davoli, R., Costosi, E., Bigi, D., dall'Olio, S., Zambonelli, P. and Russo, V. 1992. RFLP of swine transferrin gene. *Animal Genetics* **23**, Suppl 1: 67.
Davoli, R., Zambonelli, P., Bigi, D., Farnetti, E. and Russo, V. 1994. A *Hinc*II polymorphism at the porcine γ-glutamyl transpeptidase (GGT) locus. *Animal Genetics* **25**, 57.
Dezeure, F., Vaiman, M. and Chardon, P. 1993. Characterization of a polymorphic heat shock protein 70 gene in swine outside the SLA Major Histocompatibility Complex. *Biochimica Biophysica Acta* **1173**, 17-26.
Dixon, S. C., Miller, N. G. A., Carter, N. P. and Tucker, E. M. 1992. Bivariate flow cytometry of farm animal chromosomes: a potential tool for gene mapping. *Animal Genetics* **23**, 203-210.
Echard, G., Milan, D., Yerle, M., Lahbib-Mansais, Y. and Gellin, J. 1992. The gene map of the pig (Sus scrofa domestica L.): a review. *Cytogenetics and Cell Genetics* **61**, 146-151.

Ellegren, H. 1993. Abundant (A)n (T)n mononucleotide repeats in the pig genome: linkage mapping of the porcine APOB, FSA, ALOX12, PEPN and RLN loci. *Animal Genetics* **24**, 367-372.
Ellegren, H. 1993. Variable SINE 3' poly (A) sequences: an abundant class of genetic markers in the pig genome. *Mammalian Genome* **4**, 429-434.
Ellegren, H. 1994. Linkage mapping of the apolipoprotein A1 (APOA1) gene to pig chromosome 9. *Mammalian Genome* **5**, 58-59.
Ellegren, H., Chowdhary, B. P., Johansson, M. and Andersson, L. 1994. Integrating the porcine physical and linkage map using cosmid-derived markers. *Animal Genetics* **25**, 155-164.
Farnetti, E., Davoli, R., Costosi, E. and Russo, V. 1994. An *Eco*RI polymorphism at the porcine albumin (ALB) locus. *Animal Genetics* **25**, 65.
Frelat, G., Chaput, B. and Schmitz, A. 1993. Cytometrie en flux: tri de chromosomes et PCR. In: *La PCR: in procédé de réplication in vitro* Part 1. Chapter 7.1 (D. Larzul, ed.) Collection 'Le laboratoire' Editions Lavoisier. Paris pp140-154. ISBN 2-85206-882-6.
Frelat, G., Schmitz, A. and Chaput, B. 1992. New flow cytogenetics and large genome analysis: from man to swine. *Biol Cell* **76**, 41a.
Gellin, J. and Grosclaude, F., 1991. Analyse du génome des espèces d'élevage: projet d'établissement de la carte génétique du porc et des bovins. *Prod. Anim. INRA* **4**, 97-105.
Götz, K.U. and Ollivier, L. 1992. Theoretical aspects of applying sib-pair linkage tests to livestock species. *Génétique, Sélection et Evolution* **24**, 29-42.
Groenen, M. A. M. and Ruyter, D. 1993. Het varken laat zich in de (genen)kaart kijken. *Varkens* 14 April, pp. 24-26.
Groenen, M. A. M., Crooymans R. P. M. A., Ruyter, D., van Kampen, A. J. A. and van der Poel ,J. J. 1993. Gene mapping in fram animals. In: *'Developing agricultural biotechnology in the Netherlands'*, Eds D. H. Vuijk, J. J. Dekkers and H. C. van der Plas. Pudoc scientific publications, Wageningen, pp104-112.
Gu, F., Chowdhary, B. P., Johansson, M. Andersson, L. and Gustavsson, I. 1993. Mapping of the immunoglobulin γ heavy chain *(IGHG)*, cAMP-dependent protein kinase catalytic β subunit *(PRKACB)* and transition protein 2 (TNP2)genes in pigs by *in situ* hybridization. *8th North American Colloquium on Domestic Animal Cytogenetics and Gene Mapping* Guelph, Canada, p. 69.
Gu, F., Chowdhary, B. P., Johansson, M. Andersson, L. and Gustavsson, I. 1994. Localization of the *IGHG, PRKACB* and *TNP2* genes in pigs by *in situ* hybridization. *Mammalian Genome* **5**, 195-198.
Guérin, G., Duval-Iflah, Y., Bonneau, M., Bertaud, M., Guillaume, P. and Ollivier, L. 1993. Evidence for linkage between K88 ab, K88 ac intestinal receptors to Escherichia coli and transferrin loci in pigs. *Animal Genetics* **24**, 393-396.
Haley, C. S. and Archibald A. L., 1992. Porcine genome analysis. In: *'Genome Analysis'* Volume 4: *Strategies for Physical Mapping* (edited by Kay E. Davies and Shirley M. Tilghman) published by Cold Spring Harbor Laboratory Press. pp. 99-129.
Haley, C. S. and Archibald, A. L., 1992. Genetic mapping in Meishan x Large White crosses. Proceedings of the International Symposium on Chinese Pig Breeds, Harbin, China, 11-14 August 1992, (Editor: Chen Runsheng) Northeast Forestry University Press. pp. 571-581.
Haley, C.S. and Archibald, A. L., 1992. Pig gene identification — review. In: Pork Quality: Genetic and Metabolic Factors. Papers presented at an OECD workshop, Helsinki, Finland, 8-10 June 1992. (edited by Puolanne, E., Demeyer, D. I., Ruusunen, M. & Ellis, S.) CAB International. pp. 37-57.
Haley, C. S. and Archibald, A. L., 1992. The European pig gene mapping project (PiGMaP). Simposio Internacional de Porcinocultura, Sepor '92, [International Symposium on Pig Breeding], Lorca, Spain, 15-18 September 1992, ISBN 84-87154-41-7. pp 169-176.
Haley, C. S. and Knott, S. A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**, 315-324.
Haley, C. S. and Knott, S. A. 1992. Mapping quantitative traits. In: *Third Australasian Gene Mapping Workshop*. Brisbane, Australia. pp. 26 (abstract).
Haley, C. S., Knott, S. A. and Thompson, R., 1991. Mapping quantitative trait loci in farm animals. 42nd Annual Meeting of the European Association of Animal Production, Berlin, Vol 1, 81.
Haley, C., Brown, J., McQueen, H., Couperwhite, S. and Archibald, A. 1993. The pig gene mapping project (PiGMaP). *Pig News and Information* **14**, 13N-16N.

Johansson, M., Chowdhary, B. P., Gu, F., Ellegren, H., Gustavsson, I. and Andersson, L. 1993. Genetic analysis of the gene for porcine submaxillary gland mucin: physical assignment of the MUC and interferon γ genes to chromosome 5. *Journal of Heredity* **84**, 259-262.
Johansson, M., Ellegren, H. and Andersson, L. 1992. Cloning and characterization of highly polymorphic porcine microsatellites. *Journal of Heredity* **83**, 196-198.
Johansson, M., Ellegren, H., Marklund, L., Gustavsson, U., Ringmar-Cederberg, E., Andersson, K., Edfors-Lilja, I. and Andersson, L. 1992. The gene for dominant white color in the pig is closely linked to ALB and PDGFRA on chromosome 8. *Genomics* **14**, 965-969.
Johansson, M., Marklund, L., Ellegren, H., Gustavsson, U., Ringmar-Cederberg, E., Andersson, K., Edfors-Lilja, I. and Andersson, L. 1992. The gene for dominant white colour in the pig is closely linked to ALB and PDGFRA on chromosome 8. *Animal Genetics* **23**, Suppl 1:90.
Knott, S. A. and Haley, C. S. 1992. Maximum likelihood mapping of quantitative trait loci in full-sib families. *Genetics* **132**, 1211-1222.
Knott, S. A. and Haley, C. S., 1992. Interval mapping by regression — A simple method for locating quantitative trait loci in line crosses. In: *Third Australasian Gene Mapping Workshop*. Brisbane, Australia. pp. 32 (abstract).
Knott, S. A., and Haley, C. S., 1992. Aspects of maximum likelihood methods for mapping of quantitative trait loci in line crosses. *Genetical Research* **60**, 139-151.
Lahbib-Mansais, Y., Yerle, M., Dalens, M., Chevalet, C. and Gellin, J. 1993. Localization of pig Na^+, K^+ — ATPase α nd β subunit genes to chromosome 4 by radioactive *in situ* hybridization. *Genomics* **15**, 91-97.
Langford, C. F., Telenius, H., Carter, N. P., Miller, N. G. A. and Tucker, E. M. 1992. Chromsome painting using chromsome-specific probes from flow-sorted pig chromosomes. *Animal Genetics* **23**, Suppl. 1: 85.
Langford, C. F., Telenius, H., Carter, N. P., Miller, N. G. A. and Tucker, E. M. 1992. Chromsome painting using chromsome-specific probes from flow-sorted pig chromosomes. *Cytogenetics and Cell Genetics* **61**, 221-223.
Larsen, N.J. and Nielsen, V. H. 1993. *Apa*I and *Cfo*I polymorphisms at the porcine growth hormone gene. *Animal Genetics* **24**, 71.
Larsen, N. J., Nielsen, V. H. and Høj, S., 1991. Detection of a point mutation in the porcine growth hormone gene using polymerase chain reaction and restriction fragment length polymorphism. *Animal Genetics* **22**, Suppl. 1, 70-71.
Mariani, P., Johansson, M., Ellegren, H. and Andersson, L., 1991. A TaqI RFLP at the porcine platelet-derived growth factor β-receptor locus *(PDGFRB)*. *Animal Genetics* **22**, 446.
Mariani, P., Johansson, M., Ellegren, H. and Andersson, L., 1991. TaqI and PvuII RFLPs at the porcine plasminogen activator, urokinase locus (PLAU). *Animal Genetics* **22**, 447.
Mariani, P., Johansson, M., Ellegren, H., Harbitz, I., Juneja, R. K., and Andersson, L. 1992. Multiple restriction fragment length polymorphisms in the porcine calcium release channel gene *(CRC)*: assignment to the halothane *(HAL)* linkage group. *Animal Genetics* **23**, 257-262.
Mellink, C. H. M., Bosma, A. A., de Haan, N. A. and Wiegant, J., 1991. Distribution of rRNA genes in breeds of domestic pig studied by non-radioactive *in situ* hybridization and selective silver-staining. *Génétique, Sélection, Evolution* **23**, suppl. 1, 168s-172s.
Mellink, C., Lahbib-Mansais, Y., Yerle, M. and Gellin, J. 1992. Chromosomal localization of eight markers in the pig by radioactive *in situ* hybridization. *Animal Genetics* **23**, Suppl. 1:86-87.
Mellink, C., Lahbib-Mansais, Y., Yerle, M. and Gellin, J. Localization of four new markers to pig chromosomes 1, 6, and 14 by radiocative *in situ* hybridization. *Cytogenetics and Cell Genetics* **64**, 256-260.
Metezeau, P., Schmitz, A. and Frelat, G. 1993. Analysis and sorting of chromosomes by flow cytometry: new trends. *Biol Cell* **78**, 31-39.
Milan, D., Yerle, M., Schmitz, A., Chaput, B., Vaiman, M., Frelat, G. and Gellin, J. 1993. A PCR based method to amplify DNA with random primers: determining the chromosomal content of porcine flow-karyotype peaks by chromosome painting. *Cytogenetics and Cell Genetics* **62**, 139-141.
Milan, D., Yerle, M., Schmitz, A., Frelat, G. and Gellin, J. 1992. Complete characterization of porcine flow karyotype by chromosome painting after PARM-PCR. In: *'Genome Mapping and Sequencing'*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, p. 205.

Miller, J. R., Dixon, S. C. and Tucker, E. M. 1991. A pig tandem repeat sequence. *Animal Genetics* **22**, Suppl 1: 74.
Miller, J. R., Miller, N. G. A. and Tucker, E. M., 1991. Construction of a chromosome 1-specific library from the domestic pig by universal amplification of FACS purified material. Human Gene Mapping 11 (1991). *Cytogenetics and Cell Genetics* **58**, 2129-2130.
Miller, N.G.A., Langford, C.F. and Tucker, E.M. 1993. Flow cytometry of farm animal chromosomes. *17th International Congress of Genetics*, Birmingham, 15-21 August 1993, Volume of Asbstracts p. 136.
Moser, G. 1994. Analysis of influences on mapping of marker genes for quantitative traits in pig by computer simulation. Thesis, Hohenheim University, Germany.
Müller, E., Wilke, K., Moser, G. and Geldermann, H., 1991. Mapping of quantitative trait loci by means of marker genes in F2 generations of wild, Pietrain and Meishan pigs. 42nd Annual Meeting of the European Association of Animal Production, Berlin, Vol 1, 216-217.
Nielsen, V. H. and Larsen, N. J., 1991. Restriction fragment length polymorphisms at the growth hormone gene in pigs. *Animal Genetics* **22**, 291-294.
Ollivier, L. and Janss, L. L. G. 1993. A note on the estimation of the effective number of additive and dominant loci contributing to quantitative variation. *Genetics* **135**, 907-909.
Peelman, L., Van de Weghe, A., Coppieters, W., Van Zeveren, A. and Bouquet, Y. 1992. Cloning, characterization and mapping of swine lymphocyte antigen class III genes. *Animal Genetics* **23**, Suppl. 1: 40.
Peelman, L., Van de Weghe, A., Coppieters, W., Van Zeveren, A. and Bouquet, Y. 1992. Complete nucleotide sequence of a porcine HSP 70 gene. *Immunogenetics* **35**, 286-289.
Peelman, L., van de Weghe, A., Coppieters, W., van Zeveren, A. and Bouquet, Y., 1991. Cloning and sequencing of the porcine complement factor B. *Immunogenetics* **34**, 192-195.
Poulsen, P. H., Thomsen, P. D. and Olsen, J., 1991. Assignment of the porcine aminopeptidase N (PEPN) gene to chromosome 7cen-q21. *Cytogenetics and Cell Genetics* **57**, 44-46.
Russo, V., Davoli, R. and Coscelli, M. B. 1994. Identification and characterization of RFLPs in the porcine LOG12 locus. *Proceedings of the Conference on Comparative Gene Mapping in Terrestrial and Aquatic Vertebrates*, Oslo, 7-11 February 1994.
Ruyter, D., Versteghe, A. J. M., van der Poel, J. J. and Groenen, M. A. M. 1994. Five porcine polymorphic microsatellite markers. *Animal Genetics* **25**, 53.
Schmitz, A., Chaput, B., Fouchet, P., Guilly, M. N., Frelat, G. and Vaiman, M. 1992. Swine chromosomal DNA quantification by bivariate flow karyotyping and karyotype interpretation. *Cytometry* **12**, 703-710.
Schmitz, A., Chardon, P., Gainche, I., Chaput, B., Guilly, M. N., Frelat, G. and Vaiman, M. 1992. Pig standard bivariate flow karyotype and peak assignment for chromosomes X, Y, 3 and 7. *Genomics* **14**, 357-362.
Signer E.N. and Jeffreys A.J. 1993. Application of human minisatellite probes to the development of informative DNA fingerprints and the isolation of locus-specific markers in animals. In: *DNA fingerprinting: state of science*. Birkhäuser Verlag Basel, Switzerland. pp. 421-428.
Signer, E. N. 1994. Digoxigenin labelling of RNA transcripts from multi- and single-locus DNA minisatellite probes (chapter 12). In: *Protocols for nucleic acid analysis by nonradioactive probes*. Ed. P.G. Issac. Methods in Molecular Biology, vol **28**, Humana Press, New York.
Signer, E. N. 1994. Southern blot hybridization using digoxigenin-labelled RNA minisatellite probes and colour detection (chapter 18). In: *Protocols for nucleic acid analysis by nonradioactive probes*. Ed. P.G. Issac. Methods in Molecular Biology, vol **28**, Humana Press, New York.
Signer, E. N. and Jeffreys, A. J. 1992. *Fingerprint News* **4/2**: 3-7.
Thomsen, P. D., Hindkjaer, J. and Christensen, K. 1992. Assignment of a porcine male-specific DNA repeat to Y chromosomal heterochromatin. *Cytogenetics and Cell Genetics* **61**, 152-154.
Wilke, K., Jung, M., Chen, Y. and Geldermann, H. 1994. Porcine (GT)n sequences: structure and association with dispersed and tandem repeats. *Genomics* **21**, 63-70.
Winterø, A. K. and Fredholm, M. 1991. Isolation of porcine microsatellites. *Animal Genetics* **22**, Suppl 1:92.
Winterø, A. K. and Fredholm, M. 1992. Genetic variation at twenty four porcine (dG-dT)n — (dC — dA)n microsatellite loci. *Animal Genetics* **23**, Suppl. 1: 71.
Winterø, A. K. and Fredholm, M., 1991. Towards construction of a high resolution map of the porcine genome using evolutionary conserved sequences. Human Gene Mapping 11 (1991). *Cytogenetics and Cell Genetics* **58**, 2132.

Winterø, A. K., Fredholm, M. and Thomsen, P. D. 1994. A porcine polymorphic microsatellite locus (S0077) at chromosome 16q14. *Animal Genetics* **25**, 122.
Winterø, A. K., Fredholm, M. and Thomsen, P. D., 1992. Variable $(dG\text{-}dT)_n$- $(dC\text{-}dA)_n$ sequences in the porcine genome. *Genomics* **12**, 281-288.
Yerle, M., Galman, O. and Echard, G., 1991. The high resolution GTG-banding pattern of pig chromosomes. *Cytogenetics and Cell Genetics* **56**, 45-47.
Yerle, M., Galman, O., Lahbib-Mansais, Y. and Gellin, J., 1992. Localisation of the pig luteinizing hormone/choriogonadotropin receptor gene (LHCGR) by radioactive and non-radioactive *in situ* hybridization. *Cytogenetics and Cell Genetics* **59**, 48-51.
Yerle, M., Schmitz, A., Milan, D., Chaput, B., Monteagudo, L., Vaiman, M., Frelat, G. and Gellin, J. 1993. Accurate characterization of porcine flow karyotype by PCR and fluorescence *in situ* hybridization. *Genomics* **16**, 97-103.
Zambonelli, P., Davoli, R., Bigi, D., Farnetti, E. and Russo, V. 1993. A BglII polymorphisms at the porcine γ-glutayl transpeptidase (GGT) locus. *Animal Genetics* **24**, 217.
Zambonelli, P., Davoli, R., Bigi, D., Farnetti, E. and Russo, V. 1993. RFLP in porcine GGT gene. *Proceedings of the 44th Meeting of the European Association of Animal Production* Vol 1, 115, Abstract.
Zambonelli, P., Davoli, R., Dall'Olio, S and Russo, V. 1993. A MspII polymorphisms at the porcine γ-glutayl transpeptidase (GGT) locus. *Animal Genetics* **25**, 66.
Zijlstra, C., Bosma, A. A., De Naan, N. A. and Meijerink, P. H. S. 1992. Chromosomal assignment of the pepsinogen A gene in the pig by *in situ* hybridization. *Animal Genetics* **23**, Suppl. 1:84.

A study of fish genes and the regulation of their expression (BIOT CT-900188)

COORDINATOR:
F. OLLEVIER, K.U. Leuven, Leuven, B
PARTICIPANTS:
A. BELAYEW, Univ. de Liège, Liège, B
C. LECOMTE, Eurogentec S.A., Liège, B
L. COLOMBO, Univ. di Padova, Padova, I
M. CARILLO, C.S.I.C., Torre de la Sal, E
A. BRUNEL, Sepia International, Saint-Quentin-en Yvelines, F
D. CHOURROUT, INRA, Jouy-en-Josas, F
M. SCHARTL, Univ. of Würzburg, Würzburg, D
F. GANNON, Univ. College, Galway, IRL
L. PILSTRÖM. Uppsala Univ., Uppsala, S

BACKGROUND INFORMATION

In lower vertebrates, relatively few studies have addressed the **identification of key genes involved in major functions** and the **characterisation of fish gene structure and the regulation of their expression.** This lack of fundamental studies is explaining among others, the present poor understanding of the molecular biology of growth, sexual maturation, tumor formation and pathogen sensitivity. The benefits of this knowledge to aquaculture and biomedicine are expected to be large.

Aquaculture is becoming a major source of food and is slowly bridging the gap with wild catch fisheries. Consequently the significance of fish husbandry is rising. If just a few percent of the world fish production originated from controlled systems in 1950, this value has risen to 12% in 1992. Western Europe harvests about 12% of the world aquaculture production from intensive and semi-intensive salt water and fresh water production facilities. Several constraints limit the output of aquaculture. Diseases have a dramatic impact on fish in the confinement of ponds, raceways, tanks and reservoirs. Optimal nutrition, the control of sexual maturation, breeding and physical adaptation represent additional limitations.

Besides the expected applications of a better understanding of fish molecular biology in aquaculture, several developmental studies are in progress on fish models. They will reveal that biological systems are conserved in vertebrate evolution up to humans. This will raise the interest of biomedical researchers. One has to think of the regulation of oncogenes in the swordtail *(Xiphophorus)*, the immunology of bony fish and the role of homeobox genes in zebrafish development. Of special interest are the liver for smoltification and the pituitary gland which regulates juvenile development, puberty, gonadal differentiation and aging. Fish are good models because they are easy to raise, produce numerous externally fertilised eggs, have a well known developmental biology, adapt to salt and fresh water and have been the focus of fine genetic studies.

In summary, a better knowledge of fish genes is essential to support any applied biological research. Preparing all-fish inducible promoters, and linking them to fish genes of interest will be beneficiary to industry.

OBJECTIVES AND PRIMARY APPROACHES

The general objectives are the isolation and characterisation of several fish genes, including their promoter/enhancer sequences. These genes play a special role in fish biology and have a potential for applications. The characterisation occurs as well on a molecular, cellular as organism level.

The specific objectives can be grouped under the following headings:

(1) *Cloning of selected fish cDNA's encoding:*
— cod and rainbow trout immunoglobulins (Ig)
— sea bass prolactin (PRL) and growth hormone (GH)
— rainbow trout Pit-1
— liver proteins involved in smoltification of Atlantic salmon

The strategy used is the isolation of mRNA's from fish tissue, the construction of a cDNA library, the screening with different types of probes, the isolation of the selected DNA and the determination of its sequence.

(2) *Isolation of selected genes, including the promoter/enhancer sequences of:*
— rainbow trout metallothionein and growth hormone
— *Xiphophorus* oncogene
— cod and rainbow trout immunoglobulins
— silver carp growth hormone
— tilapia prolactin I and growth hormone
— rainbow trout Pit-1
— the hepatocyte nuclear factor 1, albumin and several globin genes of Atlantic salmon

The strategy involves the isolation from genomic DNA of the sequence of interest by hybridisation with a cDNA sequence. Thereafter the nucleotide sequence including the exons and introns is determined.

(3a) *Characterisation of the promoters/enhancers* in vitro:
— tilapia prolactin I and growth hormone promoter linked to the firefly luciferase gene (Luc)
— mouse immunoglobulin promoter linked to the chloramphenicol acetyltransferase gene (CAT)
— proximal promoter of the hepatocyte nuclear factor I of Atlantic salmon linked to CAT

Deletion mutants of the promoter/enhancer, linked to a reporter gene (Luc or CAT) are tested in primary/stabilised fish cells and in the absence of these in established mammalian cell lines. Methods of interest include nucleotide sequence comparison, DNase protection assays (footprinting), and electrophoretic mobility shift assays.

(3b) *Characterisation of the promoters/enhancers* in vivo:
First a methodology to transfer genes in zygotes to generate transient and stable transgenics had to be developed. This included the production of heterozygous gynogenetic progeny to duplicate the inserted hybrid gene.
The following systems were tested in fish:
— mouse immunoglobulin promoter in rainbow trout (*Oncorhynchus mykiss*)
— tilapia prolactin I promoter in African catfish (*Clarias gariepinus*)
— mammalian GH in rainbow trout

The standard procedure to test hybrid genes containing the promoter/enhancer of interest linked to a reporter gene is their microinjection in oocytes or in newly fer-

tilised eggs. Either the transient expression (occurring during embryonal development) or the stable (=heritable) expression is characterised for its time- and tissue-specificity in a backcross.

(4a) *Construction of bacterial expression vectors containing cDNA of fish origin:*
— expression of recombinant rainbow trout growth hormone and prolactin

(4b) Construction of fish expression vectors containing hybrid genes of fish origin and expressed in vitro *or* in vivo:
— rainbow trout metallothionein A promoter linked to silver carp growth hormone
— efficient expression of various heterologous vectors
— *Xmrk proto-oncogene* under control of the metallothionein promoter of rainbow trout in medaka (*Oryzias latipes*)

(5) *Sex control of sea bass by means of ploidy manipulation and hormonal treatment.* The sex ratio is manipulated with a hormonal treatment of steroids. Sterile triploid progeny are produced by temperature or pressure shocks.

(6) *A demonstration project involves the characterisation of the nucleocapsid gene of the Viral Hemorrhagic Septicemia Virus* (VHSV).

RESULTS AND DISCUSSION

(1) The cloning and sequencing of the following cDNA sequences has been realised:
— Immunoglobulin (Ig) heavy and light chain sequences were isolated from rainbow trout and Atlantic cod by antibody screening of cDNA expression libraries.
— Several cDNA's of mRNA induced during smoltification of Atlantic salmon were cloned: cytochrome c oxidase and transferrin (obtained by differential hybridisation); apolipoprotein A-1, β-fibrinogen, complement component C3 and hemopexin.
— The *hepatocyte nuclear factor 1* (HNF1) of Atlantic salmon shows 46% homology with the rat HNF1 factor; α- and non-Bohr β globins, which have a higher affinity for oxygen, were also isolated.
— The cDNA of rainbow trout *Pit-1* encodes a 358 amino acid protein. The protein shows a high affinity but low species-specificity in fish and mammals for the GH and PRL promoters.
— Sea bass *prolactin* exists in only one form which contains 188 amino acid residues. Sea bass *growth hormone* contains 187 amino acid residues and has a high degree of conservation with rtGH and rtPRL.

(2) The following fish genes including their promoter were isolated in order to characterise their sequence and intron/exon structure:
— The rainbow trout *metallothionein A* (rtMT-A) gene spans 1.1 kb, and has 3 exons and 2 introns.
— The *Xmrk oncogene* of *Xiphophorus*, located at the Tu locus, has been isolated and sequenced.
— The organisation of the rainbow trout *immunoglobulin* (Ig) *heavy chain* and the rainbow trout and Atlantic cod *Ig light* chain gene are of the clustered type. The V_L (variable), J_L (joining) and C_L (constant) segments are closely linked, separated by 0.6 kb. No coevolution of heavy and light immunoglobulin chains exists in teleost fish. The J-segment and beginning of the C_H1 segment of the rtIg chain were sequenced. V_H clones of cod have been divided in three different families, which are more conserved between species than within any species.

— The tilapia *prolactin I* gene extends over 3.7 kb and contains five exons and four introns; it has an intermediate size between those of carp and mammals.
— The *growth hormone* gene of silver carp (scGH) spans a region of 2.5 kb and consists of 5 exons. The tilapia growth hormone gene extends over 1.6 kb and contains six exons just as in salmonids but unlike the GH genes of silver carp, common carp, grass carp, mammals and birds.
— Respectively 430 and 260 bp of the rainbow trout *metallothionein A and B* promoter (rtMTa and rtMTb) were sequenced. rtMTa contains two metal responsive elements.
— 3.4 kb of the tilapia *prolactin I* proximal promoter ($tiPRLI_{3.4}$) and the tilapia growth hormone promoter ($tiGH_{3.5}$) were sequenced.
— The serum albumin promoter has a highly conserved HNF1 consensus sequence. The promoter region of HNF1 reveals a series of putative HNF4 binding sequences. α- and β globin genes are interestingly arranged in a tail-to-tail orientation.
— The proximal promoter of the rainbow trout *growth hormone* (rtGH) gene was sequenced.

(3a) Then the full characterisation of the promoter/enhancer sequences was done *in vitro*:
— *$tiPRL_{3.4}$ and $tiGH_{3.5}$*: eight and resp. fifteen putative Pit-1 transcription factor binding sites were identified of which 3 resp. 5 were confirmed by DNase I footprinting with rat nuclear extracts. Transfection experiments in established rat pituitary cell lines and fish cells by means of electroporation of deletion mutants showed a high level of expression. Cotransfection experiments of the GH or PRL promoter with a rat Pit-1 expression vector in non-pituitary fish cells (EPC) resulted in strong expression. This research was done in collaboration with the industrial partner.
— Within the proximal 466 bp of the *rtGH* promoter are located several Pit-1 binding sites and a cAMP responsive element. Rat and fish Pit-1 can activate the rtGH promoter.
— The transfactor Pit-1 shows equally important interactions on rainbow trout prolactin.
— *HNF1* of Atlantic salmon was expressed in a human hepatoma cell line: it contains negative and positive regulatory elements.
— The rainbow trout *metallothionein* a and b promoters are induced by Zn^{2+} and Cd^{2+}. The 120 bp proximal promoter is sufficient for metal induction.

(3b) The full characterisation of a fusion gene *in vivo* can be done either in a transient system or a stable transgenic system.
— *Stable transgenesis* was optimised by microinjection in medaka, rainbow trout and African catfish. It results in the mosaic presence of the hybrid gene in most founders, a transmission to 10 to 20% of the F_1 progeny and a Mendelian inheritance in F_2 progeny. The integration of the transgenes is now certain in these species. Despite a genuine effort, sperm incubation of hybrid genes in rainbow trout was unsuccessful.
— The *oncogene Xmrk* was functionally analysed in a transient system (medaka) to evaluate its tumor-inducing capacity after DNA transfer in tumor-free individuals and to study the embryonic function of the growth factor receptor encoded by the *Xmrk* proto-oncogene for cell growth and differentiation during development. Overexpression of the oncogene *Xmrk* leads to abnormal development, tumor-like lesions and early embryonic mortality. The gene is of interest in pathological processes of tumour formation in fish.

— *rtMTa-CAT* is cell-type specific in medaka and is suitable for the inducible expression of fish genes.
— The hybrid gene consisting of the *mouse Ig promoter linked to CAT* induced B-lymphocyte-specific expression in rainbow trout.
— Preliminary experiments were done with rtMTa and salmon α globin in rainbow trout.
— A detailed study of expression of the *mammalian growth hormone* gene in rainbow trout shows that this occurs in very rare cases due to the nature of the vectors. Expression was rescued in very few cases due to peculiar position effects.
— The *tilapia prolactin promoter* shows negative and positive regulatory elements when transiently expressed in embryos of African catfish during early development. Further characterisation of the prolactin promoter by means of stable integration is ongoing. Heterozygous and homozygous gynogenetic African catfish were produced by pressure and temperature shocks or their combinations. It is now possible to duplicate single site integrations of a fusion gene to homozygosity.

(4a) Important for the expression of hybrid genes is the development of efficient expression vectors in bacteria.
— Large amounts of recombinant sea bass and rainbow trout growth hormone and prolactin could be produced with a fusion gene consisting of the target protein and β-galactosidase in the *pAX4a* + vector.

(4b) Alternatively, hybrid genes can be expressed in cell lines and in embryos:
— The *rainbow trout metallothionein A promoter and silver carp growth hormone* (rtMTa-scGH) hybrid gene is active in most cell lines tested at the mRNA level.
— A whole suite of *heterologous hybrid genes* was tested to increase the level of expression: the test combined 3 kinds of enhancers/promoters; two 5' untranslated regions, 4 introns and 2 terminators and 2 reporter genes. Complex interactions among the various elements of the hybrid genes showed the importance of correct constructs. Efficient constructs are now available for all genes studied.

(5) In a project aimed at the molecular characterisation of sea bass, a strategy for *sex control* has been developed. First, the natural differentiation of the gonads was characterised histochemically. Then feminising and masculinising hormones were administered to determine the labile period of sex reversal and to generate neomales (= sex reversed females) and neofemales (= sex reversed males). To identify the sex determining mechanism, crosses of females with neomales resulted in almost exclusively male families. This project has been realised in collaboration with the industrial partner. Also, in order to sterilise sea bass, triploidisation with pressure and thermal shocks was successfully achieved.

6) In a demonstration project *viral genomes* were cloned and expressed; induction of transient but not stable expression of the nucleocapsid gene of the Viral Hemorrhagic Septicemia Virus was obtained.

In conclusion, the successful molecular characterisation of fish genes at various levels of organisation, prepared the path for various applications.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The engineering of several all-fish hybrid genes is probably the most significant achievement. Until now, expression of hybrid genes (usually consisting of

heterologous elements) in fish was all too often problematic. A *new generation of all-fish promoters and genes* has been developed. The promoters can be used for time- and tissue-specific expression and provide thus a 'casette'-type system for the inducible expression of the gene of choice.

Several fish genes were isolated and characterised. *Transient expression of gene constructs in cell culture and in vivo* was achieved. Moreover, *stable expression of hybrid genes in vivo* has been achieved. An optimised method of *expressing recombinant fish growth hormone and prolactin in bacterial cells* was developed. All these elements have a potential for industrial applications.

The small aquarium fish *(medaka)* has proven to be a very useful model for research on the regulation of gene expression. An international programme on this topic will be extremely valuable.

Finally, molecular characterisation of genes induced by the smoltification process in Atlantic salmon, the affinity of hemoglobin for oxygen, oncogenes and pituitary hormones in fish represent considerable scientific progress. Manipulation of sex control by genetic engineering in sea bass is possible.

MAJOR COOPERATIVE LINKS

Contacts among the 8 academic and 2 industrial partners have stimulated extensive collaboration. Formal collaboration included meetings, workshops, exchanges of researchers and material, publications and new grant applications. Six coordination meetings within the framework of the European Laboratories Without Walls (ELWW) were organised (Leuven, Padova, Jouy-en-Josas, Liège, Galway and Castellon). Two of these meetings received special financial support of the European Commission. There was a strong feeling among the partners that the quality of these meetings increased considerably with time. One workshop on 'Gene expression and regulation' was organised with major support of the European Union and the participation of several members of the ELWW in the teaching programme at the University of Würzburg in February 1993. Fifteen participants were selected and participated in an intensive lecturing and laboratory programme which centred around the molecular genetics of medaka. Exchanges of researchers and students were organised between several groups (an estimated total of 15 man/months). Materials exchanged among the participants included vectors, genes, promoters/enhancers, stable cell lines, sequences, fish eggs and embryos. Seven joint papers have already been published in internationally refereed journals; several more are in preparation. Finally, new international grant applications are the direct outcome of collaboration among the partners (Human Capital and Mobility, AIR, Biotech and Copernicus).

Informal collaboration includes the sharing of scientific information between academic researchers and industry, the exchange of technical and scientific knowledge, the generation of new ideas and the creation of a stimulating environment for research.

PUBLICATIONS

Joint publications

Argenton F., Vianello S., Bernardini S., Jacquemin P., Martial J.A., Belayew A., Colombo L. and Bortolussi M. (1993) The transcriptional regulation of the growth hormone gene is conserved in vertebrate evolution. Biochim. Biophys. Res. Comm. **192**: 1360-1366

Bearzotti M., Perrot E., Michard-Vanhee C., Jolivet G., Attal J., Théron M.-C., Puissant C., Grabowski H., Dréano M., Kopchick J.J., Powell R., Gannon F., Houdebine L.-M. and Chourrout D. (1992) Gene expression following transfection of fish cells. J. Biotech. **26**: 315-325

Cavari B., Funkenstein B., Chen T.T., Gonzalez-Villasenor L.I. and Schartl M. (1993). Effect of growth hormone on the growth rate of the gilthead seabream (*Sparus aurata*), and use of different constructs for the production of transgenic fish. Aquaculture **111**: 189-197
Michard-Vanhée C., Chourrout D., Strömberg S., Thuvander A. and Pilström L. (1994) Lymphocyte expression in transgenic trout by mouse immunoglobulin promoter/enhancer. Immunogenetics, **40**: 1-8
Nakayama I., Foresti F., Tewari R., Schartl M. and Chourrout D. (1994) Sex chromosome polymorphism and heterogametic males revealed by two cloned DNA probes in the ZW/ZZ fish *Leporinus elongatus*. Chromosoma **103**: 31-39
Sekkali B., Hellemans B., Martial J.A., Belayew A., Ollevier F. and Volckaert F. (1994). A comparative study of reporter gene activities in fish cells and embryos. Mol. Mar. Biol. Biotech. **3**: 30-34
Volckaert F., Hellemans B., Galbusera P., Sekkali B., Belayew A. and Ollevier F. (1994). Replication, expression and fate of foreign DNA during embryonal and larval development of the African catfish (*Clarias gariepinus*). Mol. Mar. Biol. Biotech. **3**: 57-69

INDIVIDUAL PUBLICATIONS

Adam, D., Mäueler W. and Schartl M. (1991) Transcriptional activation of the melanoma inducing Xmrk oncogene in *Xiphophorus*. Oncogene **6**: 73-80
Adam D., Dimitrijevic N. and Schartl M. (1993) Tumor suppression in *Xiphophorus* by an accidentally acquired promoter. Science **259**: 816-819
Altschmied J. and Schartl M. (1994) Genetics and molecular biology of tumor formation in Xiphophorus. In: S. Gordon (ed.) The legacy of cell fusion. Oxford University Press, Oxford
Bengtén E., Strömberg S. and Pilström L. (1994) Immunoglobulin VH regions in Atlantic cod (*Gadus morhua* L.): their diversity and relationship to VH families from other species. Dev. Comp. Immunol. **18**: 109-122
Bengtén E., Daggfeldt A. and Pilström L. (1994) The immunoglobulin light chain loci in teleost fish have an elasmobranch-like organisation. Acta Univ. Upsal. no. 18 and 19
Bernardini S., Panozzo C., Argenton F., Bortolussi M. and Colombo L. (1993) Strategies for the production of recombinant growth hormone of rainbow trout (*Oncorhynchus mykiss*). Riv. Ital. Acquacol. **28**: 117-126
Byrnes L. and Gannon F. (1992) Sequence analysis of a second cDNA for Atlantic salmon (*Salmo salar*) serum albumin. Gene **120**: 319-320
Byrnes L., Enenkel B., Smith T. and Gannon F. (1994) Preparation of nucleic acids for use as probes. Gene probes: a practical approach Vol.I, IRL Press (in press)
Carrillo M., Zanuy, S., Blazquez, M., Ramos, F., Piferrer, F. and Donaldson, E.M. (1994). Sex control and ploidy manipulation in sea bass. In: Proc OECD Symposium on environmental impacts using organisms derived through modern biotechnology. CON/ENV/DSTI/EPPC/BT, Trondheim 9-11 June 1993
Chourrout D. and Perrot E. (1992) No transgenic rainbow trout produced by fertilization with sperm incubated with linear DNA. Mol. Mar. Biol. Biotech. **1**: 282-285
Daggfeldt A., Bengtén E. and Pilström L. (1993) A cluster type organisation of the loci of the immunoglobulin light chain in Atlantic cod (*Gadus morhua* L.) and rainbow trout (*Oncorhynchus mykiss* Walbaum) indicated by nucleotide sequences of cDNA and hybridisation analysis. Immunogenetics **38**: 199-209
Deryckere F. and Gannon F. (1994) A one hour minipreparation technique for extraction of DNA binding proteins from animal tissues. Biotechniques (in press)
Doliana R., Bortolussi M. and Colombo L. (1992) Cloning and sequencing of European sea bass (*Dicentrarchus labrax*) growth hormone cDNA using polymerase chain reaction and degenerate oligonucleotides. DNA Sequence **3**: 185-189
Hardiman G., Wolff J., Peden J. and F. Gannon (1994) Isolation of the Atlantic salmon (*Salmo salar*) cytochrome c oxidase subunit III gene (coxII). J. Appl. Ichthyol. **10**: 64-68
Hardiman G., Byrnes L., Peden J., Wolff J. and F. Gannon (1994) Cloning and sequencing of an Atlantic salmon (*Salmo salar*) cDNA encoding cytochrome c oxidase subunit III and analysis of coxIII gene expression during parr-smolt transformation. Mol.Mar.Biol.Biotech. (in press)
Hardiman G., Byrnes L. and Gannon F. (1994) An analysis of highly expressed salmon liver genes during smoltification in the Atlantic salmon (*Salmo salar* L.) Mol.Mar.Biol.Biotech. (in press).

Hong Y. and Schartl M. (1992) Structure of the rainbow trout metallothionein A gene. Gene **120**: 277-279
Hong Y. and Schartl M. (1993) Sequence of the growth hormone (GH) gene from the silver carp (Hypophthalmichthys molitrix) and evolution of GH genes in vertebrates. Biochim. Biophys. Acta **1174**: 285-288
Hong Y., Winkler C., Brem G. and Schartl M. (1993) Development of a heavy metal-inducible fish specific expression vector for gene transfer *in vitro* and *in vivo*. Aquaculture **111**: 215-226
Houdebine L.M. and Chourrout D. (1992) Transgenesis in fish. Experientia **47**: 891-897
Lee M.A., Bengtén E., Daggfeldt A., Rytting A.-S. and Pilström L. (1993) Characterisation of rainbow trout cDNAs encoding a secreted and membrane-bound Ig heavy chain and the genomic intron upstream of the first constant exon. Mol. Immunol. **30**: 641-648
McEvoy T., Gannon F. and Screenan J. (1992) Gene transfer in fish: potential and practice. An. Biotech. **3**: 221-343
Schartl M. and Adam D. (1992) Molecular cloning, structural characterization, and analysis of transcription of the melanoma oncogene of *Xiphophorus*. Pigment Cell Research Suppl. **2**: 173-180
Swennen D., Poncelet A.C., Sekkali B., Rentier-Delrue F., Martial J.A. and Belayew A. (1992) Structure of the tilapia (*Oreochromis mossambicus*) prolactin I gene. DNA **11**: 673-684
Swennen D., Sekkali B., Poncelet A.C., Rentier-Delrue F., Martial J.A. and Belayew A. (1993) Molecular biology of tilapia prolactins. In: Aquaculture: fundamental and applied research (B. Lahlou and P. Vitiello, eds.). Am.Geoph.Un., Washington, USA
Tewari R., Michard-Vanhée C., Perrot E. and Chourrout D. (1992) Mendelian transmission, structure and expression of transgenes following their injection into the cytoplasm of trout eggs. Transg. Res. **1**: 250-260
Volckaert, F., Beullens K. and Ollevier F. (1992) The future of population and molecular genetics of marine organisms: capita selecta. In: The genetics of marine and estuarine organisms (ed. H. Hummel). Delta Instit. Hydrobiol. Res. Comm. **521**: 51-54
Volckaert, F., Galbusera P., Heremans K. and Ollevier F. (1993). Pressure-induction of diploid gynogenesis in the eggs of African catfish (*Clarias gariepinus*). Proc.IIIrd. Int. Meeting High Pressure Biology, pp. 73-81. Univ. Rochester Press, Rochester, USA
Volckaert F., Galbusera P., Hellemans B., Vanstaen D., Heremans K. and Ollevier F. (1994). Gynogenesis in the African catfish (*Clarias gariepinus* (Burchell, 1822)). I. Induction of meiogynogenesis with thermal and pressure shocks. Aquaculture, (in press)
Wagner A., Deryckere F., Hardiman G., Byrnes L. and F. Gannon (1993) The isolation and structure of liver and globin genes from Atlantic salmon. In: Aquaculture: fundamental and applied research (B. Lahlou and P. Vitiello, eds.). Am. Geoph. Un., Washington, USA
Wagner A., Deryckere F., McMorrow T. and F. Gannon (1994) Tail to tail orientation of the Atlantic salmon α and β globin genes. J. Mol. Evol. **38**: 28-35
Winkler C., Vielkind J.R. and Schartl M. (1991) Transient expression of foreign DNA during embryonic and larval development of the medaka fish (*Oryzias latipes*). Mol. Gen. Genet. **226**: 129-140
Winkler C., Hong Y., Wittbrodt J. and Schartl M. (1992) Analysis of heterologous and homologous promoters and enhancers in vitro and in vivo by gene transfer into Japanese medaka (*Oryzias latipes*) and *Xiphophorus*. Mol. Mar. Biol. Biotechnol **1**: 326-337
Winkler C., Wittbrodt J., Lammers R., Ullrich A. and Schartl M. (1994) Ligand-dependent tumour induction by receptor tyrosine kinases in transgenic medaka fish. Oncogene (in press)
Zanuy, S., Carrillo, M., Blazquez, M., Ramos, J., Piferrer, F. and Donaldson, E.M. (1994) Production of monosex and sterile sea bass by hormonal and genetic approaches. Act. Seanc. Biotechn. Sat. Bdx. Aquacult. ADA-EAS Edit. (in press)

Study of the *Avian Herpesvirus* genome of *Marek's disease virus* (BIOT CT-900173)

COORDINATOR:
L.J.N. ROSS / V. ZELNIK, IAH, Compton, GB
PARTICIPANTS:
M. BUBLOT, Laboratoire IFFA, Rhône Mérieux, F
E.F. KALETA, Justus Liebig Univ., Giessen, GR

BACKGROUND INFORMATION

Marek's disease is a lymphoproliferative disease of chickens caused by a *herpesvirus*, characterised by malignant transformation of helper T lymphocytes carrying CD4 markers. The disease continues to be of major economic importance to the poultry industry more than a decade after the development of the *herpesvirus* of turkey (HVT) vaccine.

Little is known about the genes of the *Marek's disease virus* (MDV), their role in tumour formation and about the antigens that are important for conferring protective immunity. The failure of HVT in controlling Marek's disease for the past 8 years has been attributed to the emergence of highly virulent strains of MDV. These strains were first isolated in the USA but have subsequently been reported in Europe and in most parts of the world. There is evidence that vaccination with a mixture of attenuated strains of MDV serotype-1, MDV serotype-2 and HVT is capable of preventing losses due to highly virulent strains of MDV. This vaccination strategy has been used commercially although the mechanism of the 'synergism' of the different serotypes is not known. However the MDV strains used for vaccination are cell-associated viruses which have to be preserved in liquid nitrogen, whereas HVT can be lyophilized and stored at 4-10°C. For these reasons, an HVT vaccine capable of protecting against highly virulent MDV is desirable.

One of the main aims of this project was to explore the possibility of engineering the expression of MDV antigens by HVT and to obtain a recombinant vaccine that would have the practical advantages of production and storage of HVT combined with the capacity to induce a more appropriate immune response to MDV antigens. In addition, it was also planned to examine the possibility of using HVT as a general vector for expressing antigens of other avian pathogens which are not adequately controlled by existing vaccines such as infectious bursal disease and Newcastle disease.

OBJECTIVES AND PRIMARY APPROACHES

(1) To identify and sequence glycoprotein genes of MDV.

(2) To identify and sequence non-essential genes of HVT.

(3) To construct recombinant HVT viruses expressing gB, gD and gH genes derived from the very virulent strain RB1B of MDV and to test the capacity of the recombinant vaccines to confer protective immunity to the RB1B strain of MDV.

(4) To construct HVT recombinants expressing the VP2 gene of infectious *bursal disease virus* (IBDV) and test the capacity of the recombinant vaccine to confer protective immunity to IBDV.

(5) To evaluate the safety of the HVT recombinants.

RESULTS AND DISCUSSION

At the time that the project was started, only the gC and gB genes of MDV had been sequenced and the technology for generating recombinant virus had not been developed. It was decided that IAH and Rhône Mérieux would investigate methods for generating recombinants and the laboratory at Giessen would evaluate their safety during the third year.

Initially the work at IAH focussed on sequencing the gH and gD genes of MDV and the thymidine kinase (TK) gene of HVT while Rhône Mérieux proceeded to sequence the ribonucleotide reductase gene and other potential insertion sites in HVT.

A. Sequencing of gD and gH genes of MDV and HVT.

The sequence of gH and gD genes of MDV and HVT have been determined at IAH. The results showed that the homologous genes in MDV and HVT are closely related on the basis of their amino acid sequence but that there were significant differences. The% amino acid identity between the MDV and HVT homologues being 56 and 49 for gH and gD respectively. The cloned genes are now available for insertion and expression in suitable vectors.

The MDV gD and gH genes have been expressed as fusion proteins using the prokaryotic expression vector pGEX-3X and antisera against the fusion proteins have been raised in rabbits. An anti-peptide serum against a predicted antigenic region of gD has also been raised. Both antisera reacted with the fusion proteins in immunoblots but failed to react with MDV-infected chick embryo fibroblasts (CEF). The apparent lack of expression of gD and gH in MDV-infected CEF is not known.

B. Sequencing of HVT Us region

A number of genes mapping in the short unique region (Us) of several alpha*herpesviruses* have been shown to be non-essential for virus replication *in vitro* and potentially useful as insertion sites. In order to identify homologous sites in HVT Us and to gain information about the organisation of HVT genes and their relationship to MDV genes mapping in Us, we decided to sequence the entire Us region (8.6 kbp) of HVT. This was a collaborative effort involving IAH and Rhône Mérieux laboratories.

Eight potential open reading frames (ORFs) were identified, 7 of which had counterparts in herpes simplex virus (HSV). The homologous proteins include Us 1, Us 2, Us 10, protein kinase (Us 3) and the glycoproteins gI, gD and gE. In addition, HVT contains one ORF which has a counterpart in MDV but is not homologous to any other known *herpesvirus* gene. It appears to be unique to Gallid *herpesvirus* 2 and 3. The% amino acid identity between proteins encoded by the Us of HVT and MDV ranged from 35 to 65, the most conserved protein being Us 2. Most of the genes were collinear with those of HSV except for Us 10 which is transposed in MDV and HVT. A characteristic feature of HVT is the fact that approximately two thirds of the gE gene is located in the inverted repeats flanking Us.

All the ORFs mapping in HVT Us have been expressed by *in vitro* transcription and translation. The observed M_r of US10, SORF3 and US2 were as predicted from the sequence but there were discrepancies between the observed and predicted M_rs of US1, protein kinase, gI, gD and gE which were largely due to post-translational and co-translational processing. These results (Zelnik et al, 1994) have provided useful information on the expected size of HVT proteins and glycoproteins and will facilitate their identification in infected cells.

C. Construction and properties of HVT recombinants

(1) Expression of MDV gB at the TK locus of HVT

We reported previously the construction of an HVT recombinant expressing the MDV gB gene at the TK locus under the control of the MDV gB promoter. Subsequent experiments at IAH have shown that the growth of the recombinant *in vivo* was impaired compared to wild type HVT, as shown by lower titres of antibodies to HVT antigens and lower numbers of infected lymphocytes. However, vaccination of a susceptible line of Rhode Island Red chickens (HPRS RIR) with 5000 p.f.u., a dose commonly used for vaccination in the field, provided total protection against challenge with the highly virulent strain RB1B of MDV. Moreover, we have also shown that vaccination with the recombinant was more effective ($P < 0.01$) than vaccination with a TK- mutant of HVT which does not contain the MDV gB gene.

These results demonstrate that HVT has potential as a vector and that MDV gB is an important immunogen which can improve the protective immune response conferred by HVT alone. However, because the wild-type HVT vaccine also provided total protection against RB1B, we have been unable to determine whether the recombinant virus is more effective than the conventional HVT vaccine. We believe that the genotype of the chickens used in our experiments is an important factor in determining their response to vaccination and might explain the unexpected efficacy of wild-type HVT against challenge with the highly virulent RB1B virus. Further experiments using different lines of chickens are warranted.

Additional evidence for the importance of MDV gB as an immunogen has been obtained using a fowlpox virus recombinant (provided by Rhône Mérieux) which expressed MDV gB. Rhode Island Red chickens vaccinated with the fowlpox recombinant at IAH were partially protected (50% surviving) whereas White Leghorns and a commercial line of chickens were better protected (75% survived, results obtained by Rhône Mérieux). These differences in the degree of protection could also be attributable to differences in the genotype of the chickens.

(2) Characterization of an HVT recombinant lacking 3 Us genes

In the process of generating an HVT recombinant in which the protein kinase gene had been targeted for insertion and expression of β-gal, we isolated a recombinant in which the LacZ gene expression cassette had been stably inserted in the short unique region (U_s) of HVT. Preliminary analysis suggested that the structure of the recombinant was not as expected. As the replication of the recombinant virus *in vitro* was not impaired, it was decided that it could be useful as a vector and that further work should be done to characterize its genome.

The results of more extensive studies involving Southern blotting, PCR analysis and sequencing revealed the following.

(a) The HVT PK gene was intact but 170 bp found at the 3'end of the gene were duplicated in the recombinant virus DNA;

(b) Three genes (US1, US10 and SORF3) present in the genome of the wild type HVT adjacent to IRs had been deleted and replaced by duplicate inverted copies of gD, gap gI and gE. The recombinant appeared therefore to contain a shorter Us and expanded IRs and TRs.

(c) The lacZ cassette was present in the opposite orientation than expected, its 3' end being closer to (IRs) than the 5' end. These results are of interest in that they suggest that expansion of the inverted repeats in *herpesviruses* might have occurred by a process of intermolecular recombination. They show moreover that three contiguous genes mapping in HVT U_s are non-essential for virus replication.

The recombinant virus was genetically stable after five passages *in vitro* and after replication in chickens. However, replication of the recombinant in vivo was impaired compared to wild-type HVT as shown by lower titres of antibodies against HVT antigens and lower numbers of infected lymphocytes. Nonetheless, the recombinant induced antibodies against β-galactosidase indicating that it has potential as a vector.

(2) Expression of IBDV VP2 gene at the ribonucleotide reductase (RR2) locus of HVT.

Experiments were carried out at Rhône Mérieux Laboratories to determine whether the small subunit of the ribonucleotide reductase (RR2) gene of HVT is a suitable locus for expression of foreign genes and whether sequences upstream of RR2 could function as a promoter to drive the expression of the IBDV VP2 gene. A transfer plasmid was constructed so that the entire RR2 open reading frame (ORF) was deleted and replaced by the VP2 ORF which started at the initiation codon of RR2, ending at the RR2 stop codon. Co-transfection of the plasmid and infectious HVT DNA resulted in homologous recombination and led to the isolation of virus recombinants which were identified by screening for expression of the VP2 gene using a monoclonal antibody. One recombinant (vHVT1) has been purified by 4 successive cycles of selection. It grows normally in chick embryo fibroblasts (CEF). Analysis of vHVT1 DNA by the polymerase chain reaction and Southern blotting confirmed the absence of RR2 sequences and showed that IBDV VP2 sequences were integrated at the RR2 locus of the HVT recombinant.

The results showed that the RR2 gene is non-essential for *in vitro* replication of HVT in CEF and that sequences upstream of the RR2 initiation codon are able to drive the expression of genes inserted into the RR2 locus. In addition, the IBDV VP2 gene expressed by the HVT recombinant appeared to be structurally identical to the native protein since it was recognised by the monoclonal antibody which is specific for conformational epitopes of VP2.

Two experiments were carried out to evaluate the efficacy of the recombinant against IBDV challenge. Day old chicks were vaccinated with different doses of the recombinant and were challenged 21 days later with the Faragher strain of IBDV. In the first experiment, birds were killed 4 days after challenge and examined for macroscopic lesions in the bursa. In the second experiment, birds were examined for gross bursal lesions 11 days after challenge and bursal/body weight ratios recorded. In the first experiment, 50% of the birds were protected by vaccination with 105 pfu. However, birds vaccinated with the recombinant were not protected in the second experiment whereas those vaccinated with a conventional inactivated vaccine were protected. Antibodies against IBDV antigens were undetectable or present in low levels in the group inoculated with the recombinant compared to those induced by the inactivated vaccine. The proportion of lymphocytes infected with

the recombinant was also significantly reduced compared to parental HVT. Moreover the degree of protection against Marek's disease was poor.

The results suggest that deletion of the small unit of ribonucleotide reductase impairs the replication of the recombinant virus and that this is responsible for the low level of protection against IBDV and Marek's disease virus.

MAJOR SCIENTIFIC BREAKTHROUGHS

The potential of HVT as a vector has been demonstrated and methods for constructing recombinants have been established. However more work is required to identify suitable sites for insertion of foreign genes without impairing the replication of the recombinants *in vivo* before the system can be exploited commercially.

The sequencing studies have shown differences in the organization of genes in the genomes of HVT and MDV, notably the absence of 3 MDV homologues in HVT Us and the fact that the gE gene of HVT, in contrast to its counterpart in MDV, is present in both IRs and TRs.

MAJOR COOPERATIVE LINKS

The IAH and Rhône Mérieux Laboratories have continued to work very closely and have exchanged immunological reagents, cloned DNAs, virus recombinants and sequence data. A notable achievement of the joint effort has been the rapid sequencing of the entire Us region of HVT. Both laboratories have benefited from the exchange of technical expertise. Meetings between IAH and Rhône Mérieux laboratories took place in Compton in June 1992, July 1994 and in Lyon in November 1991, February 1993 and October 1993. In addition, extensive discussions took place during the ELWW meetings in Denmark in November 1992 and in Tubingen in May 1994. On these occasions, Professor Kaleta and his colleagues from Giessen and Riems were also present. For administrative reasons, testing of the safety of the recombinants could not be done in Giessen as originally planned and had to be postponed. However the work is now in progress in Riems and results are expected in December 1994.

PUBLICATIONS

Individual publications

Ross, L.J.N., Binns, M.M., Tyers, P., Pastorek,J., Zelnik, V. & Scott, S.(1993). Construction and properties of a turkey *herpesvirus* recombinant expressing the Marek's disease virus homologue of glycoprotein B of herpes simplex virus. Journal of General Virology **74**, 371-377.

Scott, D., Smith, G.D., Ross, L.J.N. & Binns, M.M. (1993). Identification and sequence analysis of the gene homologues of the herpes simplex virus type 1 glycoprotein H in Marek's disease virus and the *herpesvirus* of turkeys. Journal of General Virology **74**,1185-1190.

Zelnik, V., Ross, L.J.N., Smith, G.D., Ramsay, L.A. & Pastorek, J. (1992).Comparison of glycoprotein D (gD) genes of Marek's disease virus and *herpesvirus* of turkeys. In Proceedings 19th World's Poultry Congress, Amsterdam, pp. 114-117.

Zelnik, V., Darteil, R., Audonnet, J.C., Smith, G.D., Riviere, M., Pastorek, J and Ross, L.J.N. (1993). The complete sequence and gene organisation of the short unique region of *herpesvirus* of turkeys. Journal of General Virology **74**, 2151-2162

Zelnik, V., Ross,L.J.N. and Pastorek,J.(1994). Characterization of proteins encoded by the short unique region of *herpesvirus* of turkeys by *in vitro* expression. Journal of General Virology **75**, 2747-2753.

Development of second generation vaccines against parvoviruses (BIOT CT-910256)

COORDINATOR:
J.I. CASAL, Ingenasa, Madrid, E
PARTICIPANTS:
R. MELOEN, CDI, Lelystad, NL
K. DALSGAARD, SVIV, Lindholm, DK

BACKGROUND INFORMATION

Parvoviruses cause severe disease in man and animals. *Porcine parvovirus* (PPV) is one of the major causes of reproductive failure in pigs. It causes embryonic death, mummification and stillbirths. The virus is distributed worldwide. Due to the increasing industrialization of pig farming, vaccination against disease is indispensable in order to achieve profitable meat production. The virion structure is relatively simple and consists mainly of a single protein, VP2. Canine parvovirus (CPV), a related and economically interesting parvovirus was also used as a model.

In 1991, at the beginning of the project, the development of subunit and/or synthetic vaccines based on peptides was in an impasse, due to limited success in the different approaches. But, since the emerging system of expression using *baculovirus* was highly promising, it was decided in this project to evaluate its application to parvoviruses. Also, it was decided to make a systematic approach in search of key epitopes that could be used for the design of a synthetic vaccine.

OBJECTIVES AND PRIMARY APPROACHES

To develop new generation parvovirus vaccines using efficient production procedures such as *baculovirus* expression systems and/or synthetic peptides. Additional objectives were the identification of key antigenic sites required for protection and the setting up of methods and materials for the proper formulation of vaccines.

RESULTS AND DISCUSSION

A. Procapsids vaccine

1. Production of procapsids:

PPV VP2 was cloned initially in the *baculovirus* transfer vector pJVP10Z, under the control of the polyhedrin promoter. The recombinant VP2 protein was recovered from the cytoplasm of infected Sf9 insect cells. This VP2 showed the same migration rate and immune reactivity as the viral VP2. Electron micrographs of purified preparations of recombinant VP2 showed a large number of virus-like particles (or procapsids), with a morphology very similar to the original PPV virions. The estimated amount of VP2 particles produced was between 5-10mg/10^9 cells.

2. Animal immunizations:

Prior to artificial insemination, 6 gilts were vaccinated (twice with 3 weeks interval) with a vaccine based on the recombinant protein VP2 from PPV expressed in the baculovirus/insect cell system. Four non-vaccinated pregnant gilts served as controls. The ten gilts were challenged by intravenous inoculation of a virulent strain ('839') at about 40 days of gestation and sacrificied approximately 4 weeks later.

At autopsy, the fetuses from the vaccinated gilts appeared clinically normal whereas the fetuses from the non-vaccinated gilts showed severe gross-pathological lesions, typical for intrauterine infections with PPV.

3. Effect of different constructions on the level and quality of VP2 expression.

To improve the quality of the expression, we made several new constructions using different vectors and promoters. The best choice was the vector pAcAs3, which uses the p10 promoter. In this construction, the purity and yield of the VP2 preparations was higher. Also, the preparations at the electron microscope appeared to be more homogeneous in comparison to those derived from pJVP10z.

4. Effect of the virus passage on the genetic stability.

After 20 consecutive undiluted passages of the recombinant baculovirus in insect cells, the results indicated very little effect of the virus passage on the viral titer. However, there is some effect on the total protein production and genome stability of the recombinant *baculovirus* indicating the necessity of keeping the passage of the viruses at a minimum in culture in order to have a proper virus seed for vaccine production.

B. Peptide vaccine

1. Production of specific monoclonal antibodies

Production of hybrid cells secreting monoclonal antibodies specific for PPV was performed according to well known procedures, using spleen cells from PPV-immunized mice. In summary, a total of 22 MAbs were developed.

2. Epitope mapping

Most of the mapping studies were carried out using CPV capsid proteins. Although there was some preliminary mapping studies with prokaryotic expression fragments, in this project PEPSCAN was the basic technology. Besides MAbs, a collection of dog, rabbit and guinea pig sera were used for the PEPSCAN analysis. Ten antigenic sites were located in CPV VP2. The position and surface location was determined on the 3D structure.
In PPV, PEPSCAN with polyclonal antisera from pigs and rabbits yielded 6 antigenic sites in VP2.

3. Immunization of animals

Based on these data, fourteen CPV pentadecapeptides were synthesized and used to immunize rabbits. Virus neutralizing antibodies were only obtained with two overlapping 15-mer peptides corresponding in sequence to the amino terminus of CPV VP2 (MSDGAVQPDGGQPAVRNERAT). The two peptides induced long lasting immunity (at least 8 months) either using Freund's adjuvant or aluminium hydroxide plus Quil A. A mixture of these two peptides was used to immunize dogs, giving a solid protection in all animals (10/10) against a challenge with virulent CPV. As with marker vaccines, it is possible to discriminate between vaccinated dogs not yet exposed to the virus and animals that have been infected. This is the first example of a synthetic peptide vaccine that induces real protection in target animals.
Since minks are subject to serious infection with the host range variant mink enteritis virus, a first test was done in two groups of 6 minks with the same peptide vaccine and similar baculovirus-derived CPV procapsids. All minks (12/12) were protected against highly virulent challenge.

A collection of 24 peptides corresponding to PPV antigenic sites were synthesized for rabbit immunization. From these experiments it has become clear that at least two domains of the capsid protein VP2 could be mimicked with peptides for induction of neutralizing antibodies. To further define the best peptide vaccine, a larger number of peptides have been produced and are being evaluated for neutralizing activity.

4. Optimization of peptide vaccine

An intensive search has been devoted to the selection of the amino acid sequence of the amino terminus of VP2 minimally required for induction of neutralizing antibodies (in rabbits). Experiments with over 30 different variations in sequence, length and conjugation site were carried out. Based on this information, the ultimate peptide (length, sequence, conjugation sites) for future vaccine purposes has been established.

The last and very important hypothesis to be tested is whether vaccination with the peptides and CPV procapsids in puppies can be effective even in the presence of maternally derived anti-CPV antibodies.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

A method to obtain stable parvovirus procapsids using the VP2 gene expression in insect cells has been well stablished. The structural and antigenic similarity between the parvovirus procapsids and the original virions has been demonstrated. These particles have shown a high immunogenicity, eliciting (with minimal doses of ~ 3 μg antigen/dose) an immune response in pigs (and other animals) indistinguishable from that obtained with commercial vaccines. The production of virus-like particles in the baculovirus/insect cell line has been optimized and it is already in its present form capable of producing attractive amounts of antigen at a low price, which can be used for the production of vaccines and/or diagnostic kits.

Furthermore, for the very first time a synthetic vaccine has been formulated able to confer full protection to the host animals. It is likely that this approach will be applicable for other parvoviruses, as human parvovirus B19. Optimization studies were undertaken for choosing the optimal peptide sequence, length and coupling conditions.

Both vaccines will be extremely safe (nucleic acid-free), innocuous for the animals and devoid of any residual infectivity. They should constitute the technologically improved alternative to the 'classical' vaccines.The new vaccines have been protected by patent applications.

MAJOR COOPERATIVE LINKS

Frequent meetings were realized during the project in the three countries, at least two meetings per year. The exchange of materials and information was very fluid during the whole project. Among other exchanges, C sent to A purified virus for MAb production. Several types of recombinant capsids were sent from A to C for animal experiments and E.M. studies. Synthetic peptides and immune sera were sent from B to A and C for neutralization tests. A and C prepared and sent to B a collection of MAbs, specific antisera and purified viruses for using in the PEPSCAN technology and absorption experiments.

The highly successful outcome of the project is a result of a synergistic interaction between the three participating laboratories.

PUBLICATIONS

Joint publications

C. Martinez, K. Dalsgaard, J.A.Lopez de Turiso, E.Cortés, C. Vela and J.I.Casal(1992) Production of porcine parvovirus empty capsids with high immunogenic activity. Vaccine **10**, 684-690.

J.P.M. Langeveld, J.I.Casal, C. Vela, K. Dalsgaard, S.H. Smale, W.C. Puijk and R.H. Meloen (1993) B-cell epitopes of canine parvovirus: Distribution on the primary structure and exposure on the viral surface. J. Virol. **67**:765-772.

E. Cortés, C. San Martin, J. Langeveld, R. Meloen, K. Dalsgaard, C. Vela and I. Casal (1993) Topographical analysis of canine parvovirus virions and recombinant VP2-capsids. J. Gen Virol. **74**, 2005-2010.

J.P.M. Langeveld, J.I. Casal, E. Cortés, G. Wetering, R. Boshuizen, W. Schaaper, K. Dalsgaard and R. Meloen. (1994) Effective induction of neutralizing antibodies with the amino terminus of VP2 of canine parvovirus as a synthetic peptide. Vaccine. In press.

J.P.M. Langeveld, J.I. Casal, A.D.M.E. Osterhaus, E. Cortés, R. Swart, C. Vela, K. Dalsgaard, W.C. Puijk, W. Schaaper and R. Meloen. (1994). The first peptide vaccine with real protection against viral infection in the target animal: Studies on canine parvovirus in dogs. J. Virol. **68**, 4506-4513.

J.P.M. Langeveld, J.I. Casal, C. Vela, G. Wetering, R. Boshuizen, W. Schaaper, K. Dalsgaard and R. Meloen. Delineation of optimal peptide sequence, length and coupling conditions in synthetic vaccine design. In preparation.

J.P.M. Langeveld, Uttenthal, Strandbygaard, Kamstrup,S., Dalsgaard, K., Meloen, R. and Casal, J.I.Effective vaccination of minks with synthetic peptides and recombinant VP2 capsids against mink enteritis virus. In preparation.

Kamstrup, S., Langeveld, J., Casal,J.I., Dalsgaard,K. and R.H. Meloen. Antigenic and neutralization sites in PPV mapped with synthetic peptides. In preparation.

Meloen,R., Casal, J.I., Dalsgaard,K. and Langeveld J.P.M. Prospects for peptides in vaccination against viral infections. In preparation.

Kamstrup,S., Jensen, M.H., Nielsen,J., Casal,J.I., Langeveld, J.P.M. and K. Dalsgaard.The use of antibody assay to characterize the binding of monoclonal and polyclonal antibodies to porcine parvovirus antigens. In preparation.

Kampstrup, S., Botner,A., Langeveld J.P.M., Casal, J.I. and K.Dalsgaard. The N-terminus of porcine parvovirus structural protein VP2 is exposed on the surface of parvovirus virions but not on empty capsids. In preparation.

Patents

E. Cortés, A.I.Ranz, C. Vela, K.Dalsgaard and I.Casal. Procedimiento para la producción de una vacuna subunidad contra el parvovirus porcino. P9100844(Spain, granted). EU 0554414 and USA 07-983577.

I.Casal, C.Vela, J. Langeveld, R. Meloen and K. Dalsgaard (1993) Peptidos y vacunas sintéticas contra parvovirus . PCT Patent application PCT/ES94/00006.

Expression of foot-and-mouth disease virus empty capsid particles in insect cells (BIOT CT-900190)

COORDINATOR:
G.J. BELSHAM, BBSRC Institute for Animal Health, Pirbright, Surrey, GB
PARTICIPANTS:
J.M. VLAK, Agricultural. Univ., Wageningen, NL
E. DOMINGO, CBM, Univ. Autonoma, Madrid, E
M. LOMBARD, Rhône-Mérieux, Lyon, F

BACKGROUND INFORMATION

Foot-and-mouth disease remains one of the world's most important diseases of farm animals. The disease is rapidly spread among susceptible animals which include cattle and pigs. It is caused by a *picornavirus*. Seven different serotypes of this virus have been identified, O, A, C (which are those that have been in Europe), SAT1, SAT2, SAT3 (found in southern Africa) and Asia 1 plus numerous subtypes. The disease has been successfully eliminated from Europe following intensive vaccination. The vaccine used was an inactivated preparation of virus grown in tissue culture cells. When the incidence of disease became low an increasing proportion of outbreaks became associated with the production or use of the vaccine. This contributed to the decision to cease vaccination within Europe, this will rapidly lead to the animal populations becoming highly susceptible to the disease should the virus be introduced. The potential for this to occur was demonstrated in Italy in 1993 when outbreaks of the disease were again reported. There is therefore a need for the development of vaccines against FMDV, which can be used in the face of further outbreaks, which do not need the production of large amounts of infectious foot-and-mouth disease virus.

Picornaviruses consist of a capsid with 60 copies of four different proteins, 1A (VP4), 1B (VP2), 1C (VP3) and 1D (VP1) containing a single copy of the RNA genome (about 8000 bases). A single polyprotein is encoded by the RNA which is cleaved by virus encoded proteases to produce at least 14 different polypeptides. The capsid precursor P1-2A is cleaved by the 3C protease to produce 1AB (VP0), 1C and 1D, these constitute the components of the empty capsid particle. The cleavage of 1AB to 1A and 1B is believed to occur on encapsidation of the genome. Empty capsid particles are generated in a natural infection although the level is quite different between different strains this may reflect differences in the stability of these particles or differences in the ratio of RNA to capsid protein production. The amino-terminus of the P1-2A protein is myristoylated and evidence from other picornaviruses suggests that this modification is important either for assembly or stability of the capsid.

The antigenic sites on foot-and-mouth disease virus have been mapped on a number of different strains using the selection and characterization (including sequence analysis) of mutants resistant to neutralization by monoclonal antibodies. The sites correspond to discrete clusters of amino-acids on the surface of the virus particle and involve all three surface exposed proteins (1B, 1C and 1D), hence the full spectrum of antigenicity requires the expression of the entire capsid protein precursor.

The *baculovirus* expression system is capable of expressing specific genes at high levels. The viruses are specific for particular species of dipteran cells and hence

the system is very safe for use in the production of both human and veterinary products.

OBJECTIVES AND PRIMARY APPROACHES

The primary objective was to achieve efficient expression of the FMDV structural proteins in a form in which they displayed the variety of antigenic sites exhibited by the virus particle. We anticipated that this would require the expression of the structural protein precursor with the 3C protease to permit the generation of self-assembling empty capsid particles. cDNA cassettes encoding these regions of the FMDV serotypes A, O and C were chosen for study as these represent the strains encountered in Europe previously. As mentioned above the baculovirus expression system was chosen to express these cDNAs as it represents an efficient and safe method of expressing recombinant cDNAs at high level with the advantages of appropriate protein modification machinery (eg for myristoylation). The products of these expression systems would be studied using panels of monoclonal antibodies to identify the epitopes which were presented by the recombinant proteins. Immunization experiments would test the ability of the expressed products to confer immunitiy to FMDV.

In parallel with these direct studies various approaches have been tested to enable the analyses to be performed more simply or quickly or to analyse novel features of the virus which may be exploited subsequently for its control.

RESULTS AND DISCUSSION

In the light of previous data which we had generated using the baculovirus system with FMDV cDNA we decided to express the minimal amount of the FMDV cDNAs compatible with the objective of generating a large spectrum of antigenic sites. This required mutagenesis of the cDNAs to precisely remove the coding region for the L protease (the amino-terminal portion of the polyprotein). This protease cleaves itself from the P1-2A precursor at its own carboxy-terminus and initiates the cleavage of the p220 cap-binding complex component. The removal of the L coding region made it necessary to introduce a new initiation codon. This manipulation was performed for the A10 and type C cDNAs following its success with the type O1K strain previously. These cassettes were introduced alone and with the 3C protease into baculovirus transfer vectors. Recombinant baculoviruses expressing the P1-2A cassette of type C and type A (A10) have been isolated. In each case efficient expression of the structural protein precursor was observed (3-5% of cell protein for type C and higher for the A10 cassette). The expressed product of the type C virus was characterized by Western blot analysis and using immunodot and ELISA. Suprisingly the precursor displayed reactivity with neutralizing monoclonal antibodies specific for both conformation -independent and -dependent epitopes. The reactivity with conformation dependent epitopes in both the native virus and the baculovirus expressed P1-2A precursor was lost on heating of these preparations as expected. Only weak immunogenic activity could be detected from a crude preparation of material. Analysis of the type A product is at an earlier stage but it also binds to various type A specific monoclonal antibodies. These results are important as they suggest that structural protein precursors possess structure which is similar to that of the processed and assembled molecules. It may be a very useful strategy for the development of diagnostic reagents and potentially vaccines against the wide spectrum of viruses which use the strategy of producing such a precursor. The P1-2A precursors do not assemble into macromolecular complexes as judged by sucrose density gradient analyses.

Isolation of baculoviruses containing the 3C sequence proved difficult. However, recently it has been possible for us to isolate recombinant baculoviruses which contain and express the P1-2A + 3C cassette of the A10 strain of FMDV. The correctly processed 1AB, 1C and 1D proteins have been detected and assembly into 70S material observed. This assembly is more effcient than observed previously with the O1K serotype.This data closely fits with observations made using recombinant vaccinia viruses which suggest that the A10 empty capsid particle assembles more readily than the O1K counterpart (C. Abrams et al, manuscript in preparation). Preliminary antigenic characterization of this material indicates that it displays a range of conformation-dependent epitopes shared with the native virus. Further immunogenic and structural analyses of the material will be undertaken.

The relatively poor assembly of the O1K capsid proteins into empty capsid particles we observed initially prompted us to check that the environment of the insect cell was compatible with assembly of the FMDV capsid since it is known that the FMDV particle is labile below pH6.5 and the media that are used to grow insect cells is about pH6.2. Studies were performed to determine the intracellular pH of insect cells at a range of external pH values and during a baculovirus infection. In all cases the intracellular pH, as judged by the distribution of ^{14}C -benzoic acid, was maintained close to pH7.0.

The importance of myristoylation for the assembly of FMDV empty capsids has been studied by modifying the myristoylation consensus sequence close to the amino-terminus of the P1-2A precursor. Expression of the wt and mutant P1-2A+3C cDNA cassettes was performed using the vaccinia virus T7 RNA polymerase expression system and assembly of the wt proteins into empty capsids was readily observed. However, from the mutant cassette no empty capsids were detected although efficient expression of the 1AB, 1C and 1D proteins was observed. This indicates that myristoylation is essential for FMDV empty capsid stability or assembly. Surprisingly it was also shown that the lack of myristoylation permitted the isolation of recombinant vaccinia viruses which constitutively express the P1-2A+3C(myr-) cassette whereas no recombinant vaccinia viruses have been isolated which constitutively express the native P1-2A +3C cassette.

In order to speed the analysis of the expression of cassettes within the insect cell system it was felt desirable to develop an efficient transient expression system which would enable functional assays to be performed without the requirement for generating recombinant baculoviruses. In animal cells a particularly useful system has been the introduction of the bacteriophage T7 RNA polymerease into vaccinia virus which allows efficient transient expression of plasmids containing specific gene products under the control of the T7 promoter (see above). Recombinant baculoviruses have been constructed which express the native T7 RNA polymerase or a modified version containing the SV40 Large-T antigen nuclear localization signal. In each case expression of the T7 RNA polymerase has been demonstrated and transient expression of a reporter gene (chloramphenicol acetyl transferase) and a cowpea mosaic virus 48K protein demonstrated.

In common with other picornaviruses foot-and-mouth disease virus contains an element, termed an internal ribosome entry site (IRES), within its 5' non-coding region which directs internal initiation of protein synthesis. This occurs by a cap-independent mechanism in contrast to that used by most cellular mRNAs. Mutants of foot-and-mouth disease virus with altered biological properties have been isolated following persistent infections in tissue culture. Two substitutions from the wt sequence were within the IRES. Analysis of these mutations using a bicistronic

reporter mRNA transcript revealed that one of them, at position -376 from the initiation codon, was responsible for enhanced efficiency of the mutant IRES compared to the wt sequence. This could account for the hypervirulence of the mutant virus in cells. The IRES elements have considerable utility when efficient co-expression of two gene products within mammalian cells is required.

MAJOR SCIENTIFIC BREAKTHROUGHS

Expression of the structural protein precursor P1-2A provided the first evidence that this product exhibits the complex conformation dependent epitopes displayed on the surface of the virus particle. This suggests that the surface organisation is not highly modified following cleavage by the 3C protease. This precursor may therefore be a useful reagent for diagnostic assays and potentially as a vaccine. This type of precursor is used by many viruses, not only picornaviruses, and thus this finding may have relevance to these other systems too.

The first efficient assembly of FMDV empty capsid particles by expression of cDNA in both insect and mammalian cells has been achieved. In each case correct processing of the P1-2A precursor to 1AB, 1C and 1D has been demonstrated and assembly observed using sucrose density gradients and electron microscopy.

The internal ribosome entry site (IRES) of picornaviruses is about 450 bases long and directs internal initiation of protein synthesis by a cap-independent mechanism. Characterization of this element from a mutant FMDV isolated following persistent infection in cell culture indicated that a single nucleotide substitution in the mutant was responsible for its enhanced activity and may explain the hypervirulence of this virus. This is the first demonstration of a mutant IRES displaying enhanced activity, it suggests that the virus may select for a sub-optimal IRES element, perhaps to appropriately balance the replication of the RNA with the translation rate.

MAJOR COOPERATIVE LINKS

All the studies have involved close collaboration between the different laboratories, for example the type C P1-2A precursor cDNA was prepared in Madrid, modified appropriately in Pirbright and introduced into recombinant baculovirus in Wageningen. The expressed products were analysed in different ways in all three labs. Regular meetings of participants occurred including during involvement in the ELWW on recombinant veterinary vaccines.

PUBLICATIONS

Joint publications

Martinez-Salas, E., Saiz, J-C., Davila, M., Belsham, G.J. & Domingo, E. (1993) A single nucleotide substitution in the internal ribosome entry site of foot-and-mouth disease virus leads to enhanced cap-independent translation *in vivo*. *J Virol.* **67**, 3748-3755.

Medina, M., Domingo, E., Brangwyn, J.K. & Belsham, G.J. (1993) The two species of the foot-and-mouth disease virus leader protein, expressed individually, exhibit the same activities. *Virology* **194**, 355-359.

Saiz, J-C., Cairo, J., Medina, M., Zuidema, D., Abrams, C., Belsham, G.J., Domingo,E., & Vlak, J.M. (1994) Unprocessed foot-and-mouth disease virus capsid precursor displays discontinuous epitopes involved in viral neutralization. *J. Virol.* **68**, 4557-4564.

Medina, M., Lopez-Rivas, A., Zuidema, D., Belsham, G.J., Domingo, E. and Vlak, J.M. Strong buffering capacity of insect cells: implications for the baculovirus expression system. Cytotechnology (in press).

van Poelwijk,F., Broer, R., Belsham, G.J., Goldbach, R.W. & Vlak, J.M. Development of a hybrid recombinant bacuolvirus-bacteriophage T7 transient expression system (submitted).

Individual publications

Curry, S., Abrams, C.C., Fry. E., Crowther, J.C., Belsham, G.J., Stuart, D.I. and King, A.M. (1995). Viral RNA modulates the acid sensitivity of food-and-month disease virus capsids. J. Virol. 69 (in press).

Abrams, C., King, A.M.Q. & Belsham G.J.. Expression and characterization of foot-and-mouth disease empty capsid particles from recombinant vaccinia virus vectors (in preparation)

Towards a second generation vaccine against *bovine herpesvirus* type 1 (BHV-1): immunological characterization of herpesviral glycoproteins and construction of BHV-1 deletion mutants (BIOT CT-900191)

COORDINATOR:
(1) G.M. KEIL, BFAV, Tübingen, D
PARTICIPANTS:
(2) J. van OIRSCHOT, CDI, Lelystad, NL
(3) P.P. PASTORET, Univ. of Liège, Sart-Tilman-Liège, B
(4) U.H. KOSZINOWSKI, Univ. of Heidelberg, Heidelberg, D

BACKGROUND INFORMATION

Bovine herpesvirus type 1 (BHV-1) is frequently associated with disease in the respiratory and reproductive tracts of cattle. Current vaccines do not uniformly prevent infection with virus in the field nor the establishment of latent infections. In addition, existing attenuated vaccines establish latent infections themselves. Therefore, novel more efficacious and safe BHV-1 vaccines need to be developed. With the use of biotechnology it is possible to construct negatively marked mutant virus strains with deletion(s) of one or more glycoprotein genes. However, glycoproteins of herpesviruses are a major target for the antiviral immune response of the host. In order to promote the basic understanding of the role of the immune response against glycoproteins in general and to define the potential role of the individual contribution of the different glycoproteins for vaccine preparation, studies in natural herpesvirus/host models — the infection of cattle with BHV-1 and the infection of the mouse with murine cytomegalovirus (MCMV) — were carried out.

OBJECTIVES AND PRIMARY APPROACHES

To define the contribution of individual herpesvirus glycoproteins to protective immunity as well as to herpesvirus latency. To test antigenic and immunogenic properties of viral glycoproteins of BHV-1 and MCMV following expression of individual genes coding for glycoproteins by vaccinia virus recombinants. To develop a model to test protective properties of glycoproteins with emphasis on cellular immunity *in vivo*. To delete the genes encoding gC, gE, gG and gI from the genome of BHV-1 and to assess the virulence of the respective engineered viruses. To test the mutants by challenge experiments using wild-type virus with regard to protective immunity and establishment of latency. And as long term objectives to design an engineered vaccine to BHV-1, which is immunogenic but avirulent and which does not cause latency and also to develop assays to distinguish between the antibody responses induced after vaccination and after infection.

RESULTS AND DISCUSSION

A. BFAV Tübingen

Genes coding for BHV-1 glycoproteins gG and gI were localized within the unique short segment of the viral genome. The nucleotide sequences were determined and analysis of the respective transcription units revealed that both genes are early

regulated. The open reading frames coding for gC, gE, gG and gI were integrated into vaccinia virus. The vaccinia recombinants were used to immunize rabbits to produce antisera against the corresponding BHV-1 glycoproteins. However, sera from rabbits immunized with VacgE ands VacgI did not react with proteins from BHV-1 infected cells. This may be due to the complex formation between gE and gI in BHV-1 infected celles (this result was communicated by the Lelystad laboratory). The anti VacgC reacted with BHV-1 gC and result confirmed published data obtained with monoclonal antibodies. Analysis of gG in BHV-1 infected celles and the cell culture supernatants showed that gG contains N- and O-linked carbohydrates. During intracellular transport also glucosamino glycans are added to gG which is, probably after proteolytic cleavage at the carboxy terminus to remove a putative membrane anchor sequence, subsequently secreted into the cell culture medium mainly as proteoglycan with an apparent molecular weight between approximately 90 kDa and >250 kDA. In this respect gG differs remarkably from the so far characterized gG homologs of alpha-herpesviruses. Comparison between BHV-1 and recombinant vaccinia virus expressed gG indicates that the addition of glucosaminoglycans might require an additional viral function, because this modification is not made in recombinant VacgG-1 infected cells. Further experiments demonstrated that gG is not associated with viral particles and that gG is not essential for virus replication.

BHV-1 mutants with deletions of gC, gE, gG, gI, a double deletion of gE and gI and a frameshift deletion within the gE ORF were obtained from the Lelystad laboratory and are analyzed for their virological properties in cell culture. Experiments are currently not completed.

B. ID-DLO Lelystad

1. Isotype-specific antibody responses after BHV-1 infection, reinfection and reactivation

We have developed sensitive and specific ELISAs to measure IgA, IgM, IgG1 and IgG2 antibody responses against BHV-1. Nine to 13 days after infection BHV-1-specific antibodies of the IgM isotype appeared in serum, nasal and ocular secretions and these were detectable for four weeks. The first IgA antibodies were detected a few days later than the IgM antibodies. In serum the IgA antibodies were no longer detectable after three weeks, but these did persist for prolonged periods in mucosal secretions. The calves developed a uniform IgG1 response for 13 days after infection, but the IgG2 response was quite variable; both persisted until the end of the experiment. No antibody responses were detected in genital secretions. There were no marked differences in isotype responses between calves infected with different strains of BHV-1.

After reinfection (16 weeks after primary infection) and after reactivation (22 weeks after primary infection) no BHV-1-specific IgM antibody response was detected. The serum IgA response was only transiently detectable after reinfection and again appeared rapidly after reactivation in most calves. Most calves showed an increase in nasal or ocular IgA titers after reinfection and reactivation; some calves also had IgA antibodies in genital secretions. A salient finding was that more calves showed a serum IgA response after reinfection and reactivation than virus shedding or an increase in serum IgG1 or IgG2 titer. This suggests that the serum IgA response would be the test of choice to detect BHV-1 reinfections. No correlation was found between nasal IgA titer at the time of reinfection or corticosteroid treatment and the period of virus shedding, suggesting that nasal IgA does not play a major role in protection against reinfection with BHV-1.

2. Development of marker vaccines along the classical route

We have selected a BHV-1 strain that produced small plaques *in vitro* and have shown that this strain (Za) lacks the gene coding for glycoprotein E (gE). Based on this strain a live and inactivated BHV-1 vaccine has been produced and both were found highly efficacious in vaccination-challenge experiments. In addition, an ELISA was developed to detect antibodies against gE, thus enabling differentiation between infected and vaccinated calves. These vaccines are expected to be launched on the market at the end of 1994.

3. Construction of deletion mutants of BHV-1

The gene coding for glycoprotein E was found to be located in the unique short region of the BHV-1 genome, as were the locations of the genes coding for glycoprotein I and glycoprotein G (Dr. G. Keil). The glycoprotein C (gC) maps in the unique long region. DNA fragments were constructed with deletions in either the gE, gI, gG or gC genes. By using standard transfection procedures these deletion fragments were recombined in the genome of wild-type BHV-1 strain Lam. The putative deletion mutant recombinants were selected with the use of monoclonal antibodies or glycoprotein-specific polyclonal antisera. These mutants were further examined with the Southern blot technique to verify whether they had the desired deletions. The biological properties of a gE, gI, gG and gC deletion mutant virus was evaluated in calves.

4. Evaluation of virulence, immunogenicity and latency in calves

Specific pathogen free calves were intranasally inoculated with approximately 10^5 $TCID_{50}$ of the parent wild-type strain or one of the deletion mutant strains. The control group was inoculated with virus-free culture medium. After infection, rectal temperatures and clinical observations were recorded daily for 14 days. Special attention was given to the inspection of the nasal cavity using a small flash-light. Nasal swabs were collected daily for the isolation of virus. Table 1 gives the mean ± SD number of days fever (rectal temperature >39.5°C) was recorded, virus was isolated from the nasal swabs and lesions of the nasal mucosae were observed.

Table 1: Results of first inoculation

strain	*fever (>39.5°C)*	*nasal virus shedding*	*nasal lesions*
control	0	0	0
wild-type	3.3 ± 2.3	11.2 ± 1.5	8.7 ± 1.8
gI-negative	0	8.8 ± 2.5	3.7 ± 1.8
gE-negative	0	9.8 ± 1.9	6.5 ± 2.4
gC-negative	4.8 ± 1.2	8.3 ± 1.0	8.5 ± 1.7
gG-negative	0.8 ± 0.7	10.7 ± 1.2	5.2 ± 0.8

Three weeks after first inoculation all calves were intranasally challenged with 10^7 $TCID_{50}$ of the virulent BHV-1 strain Iowa. Again for 14 days rectal temperatures were recorded, nasal swabs were collected and the nasal cavaties were observed. The results are given in table 2.

Table 2: Results of challenge

strain	*fewer (>39.5°C)*	*nasal virus shedding*	*nasal lesions*
control	6.6 ± 0.6	10.5 ± 1.2	6.0 ± 1.1
wild-type	0	0.2 ± 0.4	0.2 ± 0.4
gI-negative	0.3 ± 0.8	6.7 ± 1.5	2.5 ± 1.9
gE-negative	0	6.8 ± 1.8	0.8 ± 1.6
gC-negative	0	0.3 ± 0.5	0
gG-negative	0	2.5 ± 2.2	0.5 ± 0.8

C. University of Liège

The major aspects of the cell-mediated immune response induced by BHV-1 mutants with deletions in the genes encoding glycoproteins gI, gE, gC and gG were analysed. In this respect, peripheral blood mononuclear cells (PBMC) were isolated from the calves used in the experiment conducted at Lelystad. Proliferative assays evaluated the BHV-1 specific CD^{4+} T lymphocyte response at day 14 after inoculation with the various mutants. Positive results were observed in calves inoculated with gC-mutant of wild type BHV-1, while no significant proliferation could be detected in calves inoculated with other mutants as well as in control calves. However, experiments are still in progress to also compare proliferative responses after BHV-1 challenge in the various groups. Moreover, limiting dilution analysis will provide quantitative results allowing better comparison. CD8+ T lymphocyte cytotoxicity in PBMC isolated at day 7 after inoculation was evaluated against BHV-1 infected bovine fibroblasts. Cytotoxicity was lost after anti-CD8 monoclonal antibody and complement depletion, confirming the ability of the system to detect CD8+ activity. However, using such system, no difference could be observed between the various groups.

Stimulation of NK-like cytotoxicity was also evaluated against BHV-1 infected A549 xenogenetic tumor target cells at day 7 after inoculation. PBMC isolated from wild type BHV-1 inoculated calves had the highest cytotoxic activity. Then cells from gE-, gC and gG-inoculated calves showed slightly lower activity, followed by cells from gI-calves. Cells from the control group expressed the lowest cytotoxic activity. Since cytotoxicity was not reduced by anti-CD8 and complement depletion, the activity was thought to be mediated by CD8-cells.

In conclusion, the gC gene deletion has the lowest negative effect on the induction of the cell-mediated response. Concerning other deletions, only slight differences were observed, so that none of them appears preferable.

D. University of Heidelberg

The MCMV genes encoding the two essential glycoproteins gB and gH, glycoproteins that are conserved between most herpesviruses, were identified, sequenced, and the expressed proteins were characterized. The expected high degree of amino acid sequence homology to the respective proteins of other herpesviruses was found. Both genes were then inserted into the vaccinia virus genome and protein expression from recombinant vaccinia viruses was achieved. The posttranslational modifications of gB and gH after expression by vaccinia dif-

fered from that after expression by MCMV. For gB it was found that the precursor protein is not properly cleaved and that the protein is therefore not correctly transported to the cell membrane. Also the gH was not transported to the cell surface, but for other reasons. gH forms a heterodimer with another herpesvirus glycoprotein, this partner protein of MCMV has not yet been identified.

The experimental vaccines expressing either gB or gH were successfully tested *in vivo* for the induction of neutralizing antibodies. Antibodies raised against gB required complement for virus neutralization, whereas antibodies raised against gH were also active in absence of complement. The titers that could be achieved *in vivo* by using the vaccinia vectors were about one order of magnitude lower than the titers found after natural MCMV infection. Mice vaccinated with the vaccinia recombinants were protected against a challenge with a lethal dose of MCMV. Protection was a function of antibodies, not of T cells. Again, the protection induced by natural infection was stronger. Passive *in vivo* transfer of antisera against gB and gH resulted in protection only in the case of antibodies against gB. The immunization protocol required high doses of vaccinia and repeated immunizations, indicating that vaccinia is not an optimal vector for the intended goal.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

An attenuated and inactivated gE-negative BHV-1 vaccine was developed, in conjunction with a gE-ELISA to differentiate infected from vaccinated cattle. This vaccine will be marketed at the end of 1994. For MCMV the low efficacy of the immunization against virus challenge despite of the induction of neutralizing antibodies questioned the contribution of antibodies during natural infection. Therefore, it was studied, whether antibodies are essential for the recovery from herpesvirus infection. Mutant mice where studied that cannot produce antibodies due to the disruption of an exon of the IgM gene. It was found that antibody deficient mice recover from acute herpesvirus infection with a virus clearance kinetics that is indistinguishable form that of normal littermates which formally proves that antibodies have not an essential protective function.

MAJOR COOPERATIVE LINKS

Meetings of the project leaders and coworkers were organized every half year. Frequent additional project discussions were performed during international symposia or workshops. Direct experimental cooperation between participants from Tübingen and Lelystad to learn PCR and Tübingen and Heidelberg resulted in a joint publication. Repeated visits of a coworker from Liège to Lelystad to collect blood samples of infected cattle and to run joint experiments and to Heidelberg to learn techniques for cytotoxic T cells and antigen analysis. Exchange of information and technical help from Dr. Meloen (an ELWW lab) on the peptide specificity of antiviral antibodies. Extensive exchange of recombinant vaccinia viruses, BHV-1 deletion mutants, plasmids, antibodies, nucleotide sequences, unpublished results etc. between all participating laboratories.

PUBLICATIONS

Joint publications

M. Rapp, M. Messerle, B. Bühler, M. Tannheimer, G.M. Keil & U.H. Koszinowski. 1992. *Identification of the murine cytomegalovirus glycoprotein B gene and its expression by recombinant vaccinia virus.* J. Virol., **66**, 4399-4406.

M. Rapp, M. Messerle, B. Bühler, M. Tannheimer, G.M. Keil & U.H. Koszinowski. 1992. *Expressiopn of the MCMV glycoprotein B by recombinant vaccinia virus.* Proc. XVIIth Int. Herpesvirus workshop, Edinburgh, August 01-07, 1992, p. 281.
M. Denis, M. Slaoui, G.M. Keil, L.A. Babiuk, E. Ernst, P.-P. Pastoret & E. Thiry, 1993. *Identification of different target glycoproteins for bovine herpesvirus type 1-specific cytotoxic T lymphocytes depending on the method of in vitro stimulation.* Immunology **78**: 7-13.
M. Denis, F.A.M. Rijsewijk, J.T. Van Oirschot, P.-P. Pastoret & E. Thiry. *Identification of the role of glycoprotein gE in the cellular immune response to bovine herpesvirus 1.* Proc. 18th Int. Herpesvirus Workshop, Pittsburgh, July 25-30, 1993, p. C-12.
F.A.M. Rijsewijk, G.M. Keil, M. Maris-Veldhuis, H. Paal, R. Ruuls, A.S. Lequarré, P.-P. Pastoret & J.T. Van Oirschot. *Construction of mutants of bovine herpesvirus type 1 (BHV-1) with deletions in the nonessential glycoproteins gC, gG, gI and gE.* Proc. 3rd ELWW meeting, Lyon, F, October 7-9, 1992.
M. Denis, M.J. Kaashoek, J.T. van Oirschot, P.-P. Pastoret & E. Thiry. *Quantitative assessment of the specific CD4+ T lymphocyte proliferative response in bovine herpesvirus 1 immune cattle.* Vet. Immun. Immunopath. (in press).

Individual publications
J.T. van Oirschot, *Risk assessment of using semen from BHV-1 seropositive bulls for artificial insemination.* A. Holland Genetics Report, 1991.
J. Madic, J. Magdalena, J. Quak, M.A. Veldhuis & J.T. van Oirschot. *Epitope-specific an isotype-specific antibody responses in sera and mucosal secretions of calves experimentally infected with bovine herpes virus type 1.* Proc. 2nd Congress of the European Society for Veterinary Virology, Uppsala, 23-26 Sept., 1991, p. 129.
M.J. Kaashoek, A. Moerman, J. Madic, F.A.M. Rijsewijk, J. Quak & J.T. van Oirschot. *Evaluation of immunogenicity of a glycoprotein gE-negative strain of bovine herpesvirus type 1.* Proc. 73rd CRWAD meeting, Chicago, Nov. 9-10, 1992.
M. Denis, Slaoui, P.-P. Pastoret & E. Thiry. *Specific anti-Bovine Herpesvirus 1 cytotoxic T Lymphocytes: a method to identify their target antigen.* Proc. XVIth Int. Herpesvirus Workshop, Pacific Grove, CA, July 7-12, p. 246.
F. Fehler, J.M. Herrmann, A. Saalmüller, T.C. Mettenleiter & G.M. Keil, 1992. *Glycoprotein IV of Bovine Herpesvirus 1-expressing Cell Line Complements and Rescues an Conditionally Lethal Viral Mutant.* J. Virol. **66**: 831-839.
R.K. Maes, F.A.M. Rijsewijk & J.T. van Oirschot. *PCR based detection of bovine herpesvirus type 1 DNA in bovine semen.* Proc. XVIth Int. Herpesvirus Workshop, Pacific Grove, CA, July 7-12, p. 247.
F.A.M. Rijsewijk, J. Magdalena, J. Moedt, M.J. Kaashoek, M.A. Maris-Veldhuis, A.L.J. Gielkens & J.T. van Oirschot. *Identification and functional analysis of glycoprotein E (gE) of bovine herpesvirus type 1.* Proc. XVIIth Int. Herpesvirus Workshop, Edinburgh, August 1-7, 1992, p. 245.
F.A.M. Rijsewijk, M.J. Kaashoek, J. Magdalena, M. Maris-Veldhuis, J. Moedt, A. Moerman & J.T. van Oirschot. *Identification of the glycoprotein gE of bovine herpesvirus type 1 (BHV-1) and the vaccine properties of a BHV-1 gE deletion mutant.* Proc. 2nd ELWW meeting, Stege, DK, November 1992.
F.A.C. Van Engelenburg, M.J. Kaaashoek, F.A.M. Rijsewijk, L. Van den Burg, A. Moerman, A.L.J. Gielkens & J.T. van Oirschot. *Deletion of the glycoprotein E gene or thymidine kinase gene of bovine herpesvirus type 1 produces attenuated vaccine strains.* Proc. 17th Int. Herpesvirus Workshop, Edinburgh, August 1-7, 1992, p. 379.
F.A.M. Rijsewijk, M.J. Kaashoek, J. Madic, H. Paal, R. Ruuls, A.L.J. Gielkens & J.T. van Oirschot. *Characterization of a DNA rearrangement found in the unique short region of the Za strain of bovine herpesvirus type 1 and the vaccine properties of this strain.* Proc. 18th Int. Herpesvirus Workshop, Pittsburgh, July 25-30, 1993, p. C-67.
F.A.M. Rijsewijk, M.J. Kaashoek, A. Moerman, J. Madic & J.T. van Oirschot. *Marker vaccines and companion diagnostic test offer prospects for eradication of bovine herpesvirus type 1.* Proc. IXth Int. Congress of Virology, August 8-13, Glasgow, 1993, p. 145.
F.A.C. Van Engelenburg, R.K. Maes, J.T. van Oirschot & F.A.M., Rijsewijk. *Rapid and sensitive detection of bovine herpesvirus type 1 in bovine semen by a polymerase chain reaction based assay.* J. Clin. Microb., **31**, 3129-3135, 1993.
J.T. van Oirschot. *Conference report.* ELWW annual meeting. Vaccine, **11**, 1270-1271, 1993.

J.T. van Oirschot, M.J. Kaashoek, F.A.M. rijsewijk, A. Moerman & J. Madic. *Marker vaccines en diagnostische test bieden perspectieven voor een effectieve bestrijding van infectieuze bovine thinotracheitis.* Dier en Arts, **8**, 27-32, 1993.
J.T. van Oirschot, A.L.J. Gielkens, F.A.M. Rijsewijk, F.A.C. Van Engelenburg, M.J. Kaashoek & B. Van den Burg. *Development of a 'biotech' vaccine against bovine herpesvirus type 1 than can also serve as a vaccine vector.* In: Developing agricultural biotechnology in the Netherlands. Ed. D.H. Buijk, J.J. Dekkers & H.C. Van der Plas, Pudoc Scientific Publishers, 1993, p. 134-142.
J.T. van Oirschot, M.J. Kaashoek & F.A.M. Rijsewijk. *Do market vaccines offer prospects for eradication of bovine herpesvirus 1?* Proc. Dutch Soc. Vet. Epid. Econ., Boxtel, December 8, p. 17-30.
M. Rapp, M. Messerle, P. Lucin & U.H. Koszinowski, 1993. *IN vivo protection studies with MCMV glycoproteins gB and gH expressed by vaccinia virus.* In: Multidisciplinary approach to understanding cytomegalovirus disease. Int. Congress Series **1032**, 333-343.
M. Rapp, P. Lucin, M. Messerle, C.L. Loh & U.H. Koszinowski, 1994. *Expression of the murine cytomegalovirus glycoprotein H by recombinant vaccinia Virus. J. Gen. Virol.*, **75**, 183-188.
S. Jonjic, I. Pavic, B. Polic, L. Crnkovic & U.H. Koszinowski, 1994. *Antibodies are not essential for the resolution of primary cytomegalovirus infection but limit dissemination of recurrent virus.* J. Exp. Med., **179**, in press, May 1994.
M.J. Kaashoek, A. Moerman, J. Madic, F.A.M. Rijsewijk, J. Quak, A.L.J. Gielkens & J.T. van Oirschot. *A conventionally attenuated glycoprotein E negative strain of bovine herpesvirus type 1 is an efficacious and safe vaccine.* Vaccine, **12**, 439-444, 1994.
F.A.C. Van Engelenburg, M.J. Kaashoek, F.A.M. Rijsewijk, L. Van den Burg, A. Moerman, A.L.J. Gielkens & J.T. van Oirschot. *A glycoprotein E deletion mutant of bovine herpesvirus 1 is avirulent in calves.* J. Gen. Virol., (in press).
J.T. van Oirschot. *Vaccinations in food animal populations.* Vaccine, **12**, 415-418, 1994.
T. Engelhardt, M. Enz & G.M. Keil, 1994. *Glycoprotein G of bovine herpesvirus 1 and bovine herpesvirus 5 is a proteoglycan which is secreted from infected cells.* Proc. 3rd Congress of Veterinary Virology, Interlaken, 4-7 September 1994.

Patents

F.A.M. Rijsewijk, R.K. Maes & J.T. van Oirschot. *Bovine herpesvirus type 1 deletion mutants with natural or constructed deletion in the glycoprotein gE gene, to be used as vaccines and vaccine vectors and methods based on antibody detection against gE to distinguish infected from vaccinated cattle.* Patent pending.

AREA D:

PRENORMATIVE RESEARCH

IN VITRO EVALUATION OF THE TOXICITY AND PHARMACOLOGICAL ACTIVITY OF MOLECULES

Development of human endothelial cell lines with preserved morphological and functional characteristics. Utilization in toxicological and pharmacological tests (BIOT CT-900195)

COORDINATOR:
01. E. DEJANA, Mario Negri Inst. for Pharm. Res., Milan, IT
PARTICIPANTS:
02. J. GORDON, British Biotechnol. Ltd, Oxford, GB
03. C. de CASTELLARNAU, St. Pau Res. Fund. Biomed. Res. Inst., Barcelona, E

BACKGROUND INFORMATION

In vitro model systems of the endothelium are invaluable for understanding vascular pathology. However, the study of endothelial cells has been limited by their relative inaccessibility and short life span in culture. The need of freshly cultured endothelium also limits the possibility of an optimal standardization of functional tests. Therefore, the development of immortal endothelial cell lines would be extremely useful for the pharmacological and toxicological screenings for industrial use. For this type of application endothelial cell lines should present immortal/prolonged life span with minimal transformation i.e. retention of differentiation markers and specific functions. Retroviral and plasmid vectors containing genes able to prolong cell life have been used successfully in other cell types. This type of strategy was therefore adopted in this study.

OBJECTIVES AND PRIMARY APPROACHES

The **general objective** of this work has been to develop endothelial cell lines with indefinite/prolonged life span retaining differentiated characteristics. To reach this end point **intermediate objectives** were established:

1) to set up retroviral and plasmid vectors suitable for infection/transfection of endothelial cells.
2) to establish the best conditions for endothelial cell transfection.
3) to select the most suitable oncogenes that could prolong the life of the cells in culture without altering their differentiated characteristics.
4) to develop endothelial cell lines of murine origin from different vascular beds.
5) to develop immortalized human umbilical vein endothelial cells.
6) to characterize these cell lines for expression of specific markers and for their growth properties.
7) to set up functional tests suitable for pharmacological/toxicological screenings.
8) to apply these tests to the cell lines developed.

RESULTS

Participant 01

We succeeded in immortalizing endothelial cells with retroviral vectors containing various genes and in establishing standard functional assays for these cells.

Three immortalized **mouse endothelial cell lines** have been established. The lines were developed by infecting mouse endothelial cells or tissues with a retroviral vector containing polyoma middle T oncogene (N-TKmT, it expresses the selectable

neomycin resistance gene from the Moloney murine leukemia virus long terminal repeat and the middle T oncogene from an internal thymidine kinase gene promoter). Endothelial cell lines have been obtained from heart (H5V), brain (B9V) and whole embryo (E10V). These cells present a normal endothelial cell morphology and express von Willebrand factor. In addition they can be activated by inflammatory cytokines such as interleukin-1 to express adhesive molecules (VCAM-1, ICAM-1 and E-selectin), procoagulant activity, platelet activating factor, interleukin-6 and interleukin-1. When H5V cells were injected into immunodeficient nude mice, but not in syngeneic recipients, they developed vascular tumors. Histologically the tumours consisted of large vascular lacunae, lined by endothelial cells and a prominent mononuclear cell infiltrate. These lesions have features similar to some vascular neoplasms in man such as Kaposi sarcoma. These cells can be therefore useful in establishing in vivo models of endothelial cell pathologies. The cells have been also used successfully for the development of monoclonal antibodies in the rat directed to mouse endothelial antigens.

Two human endothelial cell lines have been developed by infection with a retroviral vector containing the oncogenes src and ras (Lsrc SN and Lras SN). In contrast to normal umbilical vein endothelial cells these cells could survive and grow in absence of growth factors. In addition they presented a higher resistance to apoptosis after detachment from the extracellular matrix in respect to normal cells. When other oncogenes were used (shc and polyoma middle T) the cells did not present immortalized characteristics but only a prolonged life span. These last cell lines could not survive in absence of growth factors.

The src and ras infected cells retained cobblestone morphology and expression of specific endothelial cell markers such as von Willebrand factor, VE-cadherin and CD31/PECAM-1. Northern blot analysis showed a good expression of cell mRNA for the infected oncogenes even at the highest cell replication number. The cell lines could be activated by inflammatory cytokines to express specific adhesive molecules such as E-selectin, ICAM-1 and VCAM-1. These cell lines were not tumorigenic in nude mice.

Participant 02

We concentrated our work on immortalization strategies with oncogenic plasmids. Two oncogenes encoding SimianVirus40 Large-Tumor-Antigen (SV40-LT-Ag) and c-myc have been selected. These oncogenes have been expressed under the control of the citomegalovirus (CMV) and Rous sarcoma virus (RSV) promoters in order to observe differences in immortalization characteristics as determined by the level of gene product expressed. The best conditions for human umbilical vein endothelial cell transfection were established (electroporation and DEAE-dextran gave the best transfection efficiency). Electroporation with the gene encoding SV40 Large T antigen under the control of the RSV long terminal repeat or the human CMV promoter produced two immortal human endothelial cell lines (DG-RSV-LT2 and DG-CMV-LT-4). These cell lines retained cobblestone morphology, endothelial cell markers (von Willebrand factor, VE-cadherin) and do not require supplementary growth factors apart from serum. In response to cytokine stimulation express adherence proteins such as ICAM-1, E-selectin and VCAM-1.

Participant 03

We first established a series of functional assays to be applyed to the endothelial cell lines developed by the other participants. In particular we standardized **assays of pharmacological interest** such as the measurement of: arachidonic metabolism,

thromboplastin, fibrinolytic activities, cell oxidation of low density lipoproteins, binding of native and either oxidized or acetylated low density lipoproteins. In addition **tests of toxicological interest** such as sensitive measurements of cell lysis and metabolic activity were also set up. These assays were applyed to the cell lines described above. In particular, this group analyzed the metabolic products of arachidonic conversion through the cycloxygenase and lipoxygenase pathways. To establish the arachidonic acid metabolic profile, the cells were incubated with different concentrations of ^{14}C-Arachidonic acid and then the culture medium was analyzed by high performance liquid chromatography. Murine cell lines synthesized mainly prostaglandin (PG) E_2 followed by HHT, PGF2α, PGD2 and HETE. In contrast human endothelial cell lines produced prostacyclin as the major arachidonic acid cycloxygenase product. This difference most probably reflects the different origin (microcirculation/ large vein) of the endothelial cell lines. Prostacyclin production was markedly increased by activation of the cells with inflammatory cytokines (IL-1 and TNF) and endotoxin.

Thromboplastin (procoagulant activity) was tested by the recalcification time assay. Endothelial cell lines produced very low amounts of thromboplastin in resting conditions but this parameter was significantly increased by cell activation with cytokines (IL-1 and TNF) and endotoxin.

In the cell lines the synthesis and secretion of plasminogen activators (tissue type, t-PA and urinary type, u-PA) and inhibitors (PAI-1) have been characterized. Essentially all the human endothelial cell lines analyzed were able to produce activators and inhibitors of plasminogen (as above) and could be activated by cytokines (IL-1 and TNF) to express higher amounts of these substances. Some quantitative differences in the production of the different agents were noticed comparing the different lines.

MAJOR SCIENTIFIC BREAKTHROUGHS

1) Establishment of murine and human immortal endothelial cell lines. These lines are now available to the scientific/industrial community.

2) Optimization of human endothelial cell infection/transfection with retroviral and plasmidic vectors. These data can have important application for gene transfer in endothelial cells.

3) Identification of oncogenes (src, ras, SV40, middle T) able to immortalize/ prolong the life span of endothelial cells. This allows to apply the same strategy for immortalizing human endothelial cells from different and more specialized vascular beds (i.e. liver, brain, lymphnodes...).

4) Standardization of functional assays of pharmacological and toxicological interest and their application to murine and human endothelial cell lines. These results constitute the starting point for the development of standardized assays for the screening of pharmacological and toxicological agents in industries.

COOPERATIVE LINKS

Seven meetings among participants along the time period of the project. One three months stage of participant 03 in participant 01 lab.. Endothelial cell lines and materials/assays have been continuously exchanged.

PUBLICATIONS

Joint publications

Lampugnani M.G., Resnati M., Raiteri M., Pigott R., Pisacane A., Houen G., Ruco L.P. and Dejana E.: A novel endothelial-specific membrane protein is a marker of cell-cell contacts. J. Cell Biol. **118**: 1511-1522, 1992.

Lopez S., Vila L., Breviario F. and de Castellarnau C.: Interleukin-1 increases 15-hydroxy-eicosatetraeinoic acid formation in cultured human endothelial cells. Biochem. Biophys. Acta **1170**: 17-24, 1993.

Re F., Zanetti A., Sironi M., Polentarutti N., Lanfrancone L., Dejana E. and Colotta F.: Inhibition of anchorage dependent cell spreading triggers apoptosis in cultured human endothelial cells. J. Cell Biol., (in press)

Individual publications

Pelicci G., Lanfrancone L., Grignani F., McGlade J., Cavallo F., Forni G, Nicoletti I., Grignani F., Pawson T. and Pelicci P.G.: A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. Cell **70**: 93-104, 1992.

Lanfrancone L., Boraschi D., Ghiara P., Falini B., Grignani F., Peri G., Mantovani A. and Pelicci P.G.: Human peritoneal mesothelial cells produce many cytokines (granulocyte colony stimulating factor (CSF) granulocyte-monocyte-CSF, macrophage CSF, interleukin-1(IL-1) and IL-6 and are activated and stimulated to grow by IL-1. Blood **80**: 2835-2842, 1992.

Dejana E.: Endothelial cell adhesive receptors. J. Cardiovascular Pharmacol. **21** (Suppl. 1): S18-S21, 1993.

Pich I., Lopez S., Vila L., Lagunas C. and de Castellarnau C.: Influence of fatty acid anilides present in toxic oils on the metabolism of exogenous arachidonic acid in cultured human endothelial cells. Toxicology **77**: 51-63, 1993.

de Castellarnau C., Pich I., Chanquia C., Vila L., Lagunas C., Fontcuberta J. and Rutllat M.: Effects of linoleic and oleic acid anilides on prostacyclin synthesis and fibrinolytic profile of human endothelial cells in culture. Relevance to the toxic oil syndrome. Toxicology **81**: 181-194, 1993.

Felez J., Chanquia C., Fabregas P., Plow E.F. and Miles L.A.: Competition between plasminogen and tissue plasminogen activator for cellular binding sites. Blood **28**: 2433-2441, 1993.

Lampugnani M.G., Resnati M., Raiteri M. and Dejana E.: Adhesion molecules at endothelial cell to cell junctions. In Cell Adhesion Molecules: Cellular adhesion mechanisms. M.E. Hemler and E. Mihich Eds. pp 149-153, Plenum Press New York, 1993.

Lampugnani M.G., Caveda L., Breviario F., Del Maschio A. and Dejana E.: Endothelial cell to cell junctions. Structural characteristics and functional role in the regulation of vascular permeability and leukocyte extravasation. In Blood Constituents and the Vessel Wall. J. Caen and G. Tobelem eds. Bailliere's Clinical Haematology Vol. 6, No. 3 pp 539-558, Bailliere Tindall, London, 1993.

Garlanda C., Parravicini C., Sironi M., DeRossi M., Bussolino F., Colotta F., Mantovani A. and Vecchi A.: Progressive growth in immunodeficient mice and host cell recruitment by mouse endothelial cell transformed by polyoma middle T: implications for the pathogenesis of opportunistic vascular tumors. Proc. Natl. Acad. Sci. (in press).

Vecchi A., Garlanda C., Lampugnani M.G., Resnati M., Stoppacciaro A., Ruco L., A. Mantovani and Dejana E.: Monoclonal antibodies specific for endothelial cells of mouse blood vessels. Eur. J. Cell Biol., (in press).

Green D.R., Banuls M.P., Gearing A.J.H., Needham L.A., White M.R.H. and Clements J.M.: Generation of human umbilical vein endothelial cell lines which maintain their differentiated phenotype. Endothelium, (in press).

Zanetti A., Conforti G., Hess S., Martin-Padura I., Ghibaudi E., Preissner K.T. and Dejana E.: Clustering of vitronectin and RGD peptides on microspheres leads to engagement of integrins on the luminal aspect of endothelial cell membrane. Blood, (in press).

Caveda L., Corada M., Martin-Padura I., Del Maschio A., Breviario F., Lampugnani M.G. and Dejana E.: Structural characteristics and functional role of endothelial cell to cell junctions. Endothelium, (in press).

Martin-Padura I., de Castellarnau C., Uccini S., Pilozzi E., Natali P.G., Nicotra M.R., Ughi F., Azzilini C., Dejana E. and Ruco L.: Expression of VE (vascular endothelial)-cadherin and other endothelial specific markers in hemangiomas. J. Pathol. (in press).

Breviario F., Caveda L., Corada M., Martin-Padura I., Golay J., Introna M., Lampugnani M.G. and Dejana E.: Molecular and functional properties of human VE-cadherin (7B4/cadherin-5) a novel endothelial specific cadherin. Blood, (submitted).

Development of *in vitro* neural systems for the identification of agents with toxicological and pharmacological potential (BIOT CT-900183)

COORDINATOR:
A: D.C. WILLIAMS, Trinity College (TC), Dublin, IRL
PARTICIPANTS:
B: A. SCHOUSBOE, Royal Danish School of Pharmacy (RDSP), København, DK
C. C. MEREDITH, Toxicology International (BIBRA), Carlshalton, GB
D: E. RODRÍGUEZ-FARRÉ, Consejo Superior de Investigaciones (CSIC), Barcelona, E
E: R. GRIFFITHS, Division of Biochemistry (USTA), St. Andrews, GB

BACKGROUND INFORMATION

In vitro systems are being used to study the mode of actions of drugs and xenobiotics as well as the structure and mechanism of the targets for these agents. At present the development of new brain-acting drugs is hindered by the lack of good *in vitro* systems to explore their mechanism of action, to screen for drug specificity and potency, and to explain and minimise toxicity and side effects. In vitro studies on human brain presently pose a difficult ethical problem in that primary cultures of central nervous system neurons and glia involve use of foetal material, a procedure which in some Community countries is socially and legally unacceptable.

OBJECTIVES AND PRIMARY APPROACHES

The aim of the project is to show the usefulness of the well established and characterised rodent primary cultures of neurons (glutamatergic and GABAergic) and astrocytes, rat PC12 cells and brain slices in *in vitro* neuropharmacotoxicology. Convulsive and anticonvulsive agents have been chosen as the compounds of interest.

RESULTS AND DISCUSSION

A. Trinity College Dublin has demonstrated previously under BAP and now in BRIDGE that peripheral-type benzodiazepines (PBs) are not antiproliferative through the PB receptor (PBR). Their anti-proliferative action has been shown not to be as previously suspected through inhibition of mitochondrial respiration. The differentiative actions of PBs has been further assessed by studying their induction of PC12 cell c-fos mRNA. Benzodiazepines such as diazepam and Ro 5-4864 can induce c-fos mRNA in addition to their known role in the superinduction of c-fos with NGF (nerve growth factor) Binding studies with astrocytic PBRs have been shown to be invalid as a screen for convulsant and anticonvulsant compounds. However identification has been made of novel compounds which stimulate and inhibit steroidogenesis possibly through PBRs. Lindane (γ-HCH) has been shown to be a stereoselective ligand for PBRs. However the lack of potency and stereoselectivity of HCH isomers in inhibiting steroidogenesis suggests that convulsive activity is not associated with the PBR. These studies report for the first time the influence of cytokines on PBR expression and their significance for illness

and toxicity. The studies also report for the first time the presence and expression of transcription factor, NFkB, in brain cells.

B. BIBRA has established a model *in vitro* system for neuronal differentiation to validate alterations in proto-oncogene expression as early markers of pre-differentiative and anti-proliferative events. In this context the rat phaeochromocytoma cell line PC12 has been established at BIBRA. The cell line has been characterised with respect to proto-oncogene expression, (c-fos, c-myc and Ha-ras) and expression of neuronal enzymes in response to Nerve Growth factor. A sensitive RT-PCR assay for the analysis of low level c-fos mRNA expression has been developed. Application of this assay has led to the following results:

a) Sustained expression of c-fos may correlate with SAA-induced toxicity in primary neuronal cultures.
b) The cytokine, IL-6 activates c-fos expression and differentiation in PC12 cells while IL-1 is produced after NGF-induced differentiation.
c) Certain benzodiazepines e.g. diazepam and Ro 5-4864 can induce c-fos mRNA in PC12 cells.

C. USTA have further developed *in vitro* primary cultures of neurons and end points such as fluorescence-based assays for intracellular free calcium levels, LSC-based assays for receptor-mediated phosphoinositide metabolism, HPLC-based assays for neurotransmitter uptake, cell monolayer continuous superfusion assays for neurotransmitter release for screening for compounds which, being excitatory amino acid antagonists, are cytotoxo-protectants and, being agonists, are cytotoxic. Primary cultures of cerebral cortex neurons and cerebellar granule cells have been successfully employed to demonstrate the cytotoxic actions of a number of endogenous excitatory amino acids directed implicated in human neuropathologies. Morphological examination and measurement of the release of cytoplasmic lactate dehydrogenase were shown to be efficient end-points for monitoring toxicity. The effects of these excitatory amino acids on the second messengers cGMP and intracellular Ca^{2+} levels (used as potential end-points for monitoring toxicity) were also fully investigated in these model *in vitro* systems. Whilst marked changes in the levels of these agents could be demonstrated, it is currently concluded that the absence of a direct correlation between toxicity and excitatory amino acid-induced increases in these second messengers precludes any single second messenger end-point being used as an unequivocal predictor of excitotoxicity. The ability of the same selective antagonists for certain excitatory amino acid receptor subtypes to protect against neurotoxicity in both cortical neurons and cerebellar granule cells by the same excitatory amino acids suggests that cultured neurons may be a useful *in vitro* system to be used in conjunction with *in vivo* tests in assessing the value of certain classes of potential neuroprotectants. The use of primary cultures of cerebellar granule cells to monitor directly formation of nitric oxide has been unsuccessful; however, a suitable and acceptable, although indirect, end-point for measurement of nitric oxide production may be the radioimmunoassay detection of cGMP levels. Further studies are needed. Excitatory amino acids of neuropathological interest induce a sustained expression of c-fos mRNA in primary cultures of cerebellar granule cells when used at high levels *in vitro*; whereas a transient expression is observed using low doses of these agents. The concentrations of these excitatory amino acids which give rise to sustained c-fos expression correlate closely to those giving rise to an increase in cytoplasmic lactate dehydrogenase release. These results suggest that induction of certain oncogenes (e.g. c-fos) may be a possible biomarker for excitotoxicity.

D. Royal Danish School of Pharmacy have further developed primary cultures of neurons and astrocytes for testing GABA uptake inhibitors as anticonvulsants, for elucidation of the differentiative actions of GABA, and similar compounds, for identification of glutamate receptor antagonists, (neurodegenerative protectants) or agonists (neurotoxins or early differentiation enhancers). The neurodifferentiative activity of GABA has been further investigated by showing: an increase in the number of voltage gated Ca^{2+}-channels and also $GABA_B$-receptors after exposure of cultured neurons to a $GABA_A$-receptor agonist. It has been demonstrated that the induction by the $GABA_A$-receptor agonist THIP of low affinity $GABA_A$-receptors is dependent on an intact polyamine level in the neurons. The *in vitro* characterisation of GABA-transporters in cultured astrocytes has led to the isolation of a neuronally derived protein which induces astrocytic GABA-transporters and *in vivo* validation of the identification of anticonvulsants from *in vitro* inhibition studies of GABA transport in cultured neurons and astrocytes. The cytotoxic action of excitatory amino acids in cultured neurons has been extensively characterised and it has been shown to involve activation of different pools of Ca^{2+} including influx and release from intracellular stores depending upon the excitatory amino aid; to involve changes in intracellular levels of cGMP presumably via a NO dependent mechanism; that selective excitatory amino acid receptor subtype antagonists can protect against neurotoxicity; that compounds affecting intracellular Ca^{2+} levels may protect against excitotoxicity and, that cultured neurons can be utilised to predict the neuroprotective value *in vivo* of agents protecting against excitotoxicity. Using an *in vitro* test system consisting of cultured neocortical neurons it has been shown, that the insecticide Lindane interacts with $GABA_A$-receptors, and that Lindane causes cytotoxic neuronal death which may be prevented by GABA and benzodiazepines.

E. Consejo Superior de Investigaciones Científicas have improved methods to study the interaction of convulsant and non-convulsant insecticides with glutamatergic and GABAergic neuronal systems using primary cultures of neurons and brain slices. Primary neuronal cultures (neocortical neurons and cerebellar granule cells) and regional brain slices (cortex and hippocampus) have been successfully employed to assess specific neurotoxic/pharmacological activity of convulsant-anticonvulsant agents (GABAergic and polychlorocycloalkanes (PCCAs): cyclodienes and hexachlorocyclohexane (HCH) isomers) using radiometric assays for neuronal $GABA_A$ receptor mediated Cl-uptake, receptor-activated phosphoinositide release, neurotransmitter (NA) release and ligand binding studies ([^{35}S]TBPS), fluorescent Ca^{2+} measurements and cytotoxicity measured by LDH release.

Convulsant and non-convulsant HCH isomers showed differential interactive effects with neurotransmitter synaptic functions: NA release, GABA-induced chloride flux, glutamate and carbachol-induced inositol phosphate formation.

At the pre-synaptic level, the convulsant γ-HCH isomer increased and the depressant δ-HCH isomer decreased noradrenaline release. At the post-synaptic level, γ-HCH inhibited and δ-HCH potentiated the GABA-induced Cl- flux. Both pre- and post-synaptic effects agree with the neurotoxic action of these compounds.

Brain slices and neuronal primary cultures have proved to be useful *in vitro* systems to provide information on the effects and mechanisms of action of convulsant and anticonvulsant agents. Noradrenaline release from brain slices may be a useful model to test the action of neuroactive and neurotoxic agents at neuronal

presynaptic level. GABA-induced ^{36}Cl-uptake in primary neuronal cultures appears to be a suitable assay to investigate the action of convulsant/anticonvulsant agents at the $GABA_A$ receptor.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

- Screening for potential anticonvulsants: primary cultures of neurons and astrocytes can be used to screen for anticonvulsant drugs; high affinity GABA uptake inhibitors can be shown in cells to protect animals against seizures; especially those selective for astrocytic uptake. **Proposed test** — easy, reliable, inexpensive, microscale and multiwell.

- Cytotoxicity test in cultured neurons — release of cytoplasmic lactate dehydrogenase; permanent damage in cerebral cortical neurons found at 55-60% release; useful for EAA cytotoxicity studies.

- Excitotoxicity tests: No single second messenger end-point assay appears suitable as predictor of EAA-mediated excitotoxicity; no simple correlation exists between toxicity and EAA-stimulated increases in a variety of second messengers. Demonstration of sustained c-fos expression in primary neuronal cultures may serve as a valid end-point for excitotoxicity: A correlation is indicated between excitotoxicity and sustained c-fos mRNA expression in cerebellar granule cells; further work is needed.

- Developmental neurotoxicity: modulation of neurite outgrowth in NGF-treated PC12 cells; preliminary evidence suggests that some developmental neurotoxins, e.g. ethanol, can modify neurite outgrowth.

- PC12 system has potential for developing both neurotoxicity and developmental neurotoxicity assays.

- Astrocytic BZ receptors shown to be invalid as screen for convulsant and anti-convulsant compounds.

- Regional brain slices and neuronal primary cultures are useful *in vitro* systems for elucidation of the pharmacotoxicology of convulsant agents.

- Noradrenaline release from brain slices may be a useful system for screening presynaptic actions of convulsants and anti-convulsants.

- GABA-induced chloride uptake in primary neuronal cultures may be a suitable assay for screening the action of convulsants and anti-convulsants at the GABA receptor.

Some work has been transferred to industry for screening tests, for application to new technology, for pharmaceutical screening of new compounds. Interaction with four companies have now been fruitful.

MAJOR COOPERATIVE LINKS

11 Project Leaders Meetings.
St. Andrews, May 1991; Puerto de la Cruz, February 1992; Dublin, August, 1992; Glasgow, October, 1992; Dublin, December 1992; London, December 1993; Ispra, February 1994; Innsbruck, April 1994, Copenhagen, May 1994.

Exchanges of personnel to:

USTA	December 1993	2 workers from RDSP — planning
	March 1992	1 Worker from TCD (7 days) — training
	October 1992	1 worker from RDSP — exchange of techniques
	March 1993	2 workers from BIBRA (4 days) — joint experiments
	October 1991	1 worker from RDSP (7 days) — technology set up
RDSP	February 199	12 workers from USTA (14 days) — cell culture training
	May 1992	2 workers from CSIC — training
	May 1992	1 worker from TCD (4 days)- joint experiments
	June 1991	1 worker from USTA (6 days) — manuscript preparation
	November 1991	1 worker from USTA (4 days) — manuscript preparation
	May 1993	2 workers from USTA — joint experiments
TCD	August 1992	1 worker from USTA (4 days)
	August 1992	1 worker from RDSP (4 days) — planning
CSIC	November 1993	2 workers from RDSP — planning
	February 1993	1 worker from TCD (30 days) — experiments
BIBRA		1 worker from USTA — training 1 worker from TCD — training 1 worker from RDSP- training

Materials and Technology Transfer

RDSP	USTA — freeze-dried samples TCD — training CSIC — training BIBRA — mRNA sample
USTA	BIBRA — SAAs RDSP — SAAs
TCD	BIBRA — cytokines, benzodiazepine CSIC — ben-zodiazepines RDSP — antisera
BIBR	AUSTA — mRNA solutions USTA — RT-PCR technology
CSIC	RDSP — compounds

Exchanges of personnel to: USTA (8), RDSP (8), TCD (2), CSIC (3), BIBRA (3).
Exchange of Materials and Technology from: RDSP (4), USTA (2), TCD (4), BIBRA (2), CSIC (1).
1 Sectoral Meeting organised (Dublin, December, 1992)

PUBLICATIONS

Joint publications

Atterwill, C.K., Bruinink, A., Drejer, J., Duarte, E., Mcfarlen Abdulla, E., Meredith, C., Nicotera, P., Regan, C., Rodríguez-Farré, E., Simpson, M., Smith, R., Veronesi, B., Vijverberg, H., Walum, E. & Williams, D.C. Report of ECVAM Neurotoxicity Working Party 1994. *ATLA Report and Recommendations of the ECVAM Workshop* on *in vitro Neurotoxicity Testing* (in press) 1994.

Dunlop, J., Grieve, A., Damgaard, I., Schousboe, A. & Griffiths, R. Sulphur-containing excitatory amino acid-evoked Ca^{2+}-independent release of D-[^{3}H]aspartate from cultured cerebellar granule cells: the role of glutamate receptor activation coupled to reversal of the acidic plasma membrane carrier. *Neuroscience* **50**, 107-115 (1992).

Dunlop, J., Grieve, A., Schousboe. A. & Griffiths, R. Stimulation of [^{3}H]GABA release from cultured mouse cerebral cortex neurons by sulphur-containing excitatory amino acid transmitter candidates: receptor activation mediates two distinct mechanisms of release. *J. Neurochem.* **57**, 1388-1397 (1991)

Frandsen, Aa., Griffiths, R. & Schousboe, A. Cytotoxic actions and effects on intracellular Ca^{2+} and cGMP concentrations of sulphur containing excitatory amino acids in cultured cerebral cortical neurons. *J. Neurosci. Res.* **34**, 331-339 (1993).

Gorman, A., Scott, M., Meredith, C. & Griffiths, R. Proto-oncogene expression in primary neuronal cultures as a possible marker for excitotoxicity. *Brain Research Association Abstracts* **11**, 42 (1994)

Gorman, A., Scott, M., Rumsby, P., Griffiths, R. & Meredith, C. Excitatory amino acid-induced neurotoxicity is associated with a sustained increase in c-fos mRNA levels. *Proceedings of CEC Biotechnology Programme (BRIDGE)* In vitro *pharmacotoxicology Sectoral Meeting, Innsbruck* (1994)

Grieve, A., Dunlop, J., Schousboe, A. & Griffiths, R. Kinetic characterization of sulphur-containing excitatory amino aciduptake in primary cultures of neurons and astrocytes. *Neurochem. Int.* **19**, 467-474 (1991).

Griffiths, R., Grieve, A., Frandsen, Aa. & Schousboe, A. Mixed actions of TMB-8 as a Ca^{2+} antagonist in cultured mouse cortical neurons. *NeuroReport* **5**, 605-608, 1994.

Griffiths, R., Meredith, C., Rodríguez-Farré, E., Schousboe, A. & Williams, D.C. Development of *in vitro* neural systems for the identification of agents with toxicological and pharmacological potentials. *Proceedings of CEC Biotechnology Programme (BRIDGE) In vitro pharmacotoxicology Sectoral Meeting, Innsbruck* (1994)

Moynagh, P.N., Zisterer, D., Schousboe, A. & Williams, D.C. Astrocyte Benzodiazepine 'Receptors': Lack of Correlation between Convulsant: Anticonvulsant Potencies and 'Receptor' Affinities. *Proceedings of CEC BRIDGE Biotechnology Sectoral Meeting on In-Vitro Pharmacotoxicology, Dublin, December 1992.*

Moynagh, P.N., Schousboe, A. & Williams, D.C. The peripheral-type benzodiazepine receptor is present in astrocytes but is not a primary site of action for convulsants/anticonvulsants. J. Neurochem. **62**, 673-679 (1994).

Moynagh, P.N., Schousboe, A. & Williams, D.C. The peripheral-type benzodiazepine receptor is present in astrocytes but is not a primary site of action for convulsants/anticonvulsants *J. Neurochem.* in press (1994)

Pomés, A., Frandsen, Aa., Suñol, C., Sanfeliu, C., Rodríguez-Farré, E. & Schousboe, A. Lindane cytotoxicity in cultured neocortical neurons is ameliorated by GABA and flunitrazepam. *J. Neurosci. Res.* (in press) 1994.

Schousboe, A., Frandsen, Aa. & Griffiths, R. The Use of Cultured Neurons to Study Involvements of Aberrations in Intracellular Ca^{2+} Homeostasis in Excitatory Amino Acid Cytotoxicity. *Proceedings of CEC Biotechnology Sectoral Meeting on In-Vitro Pharmacotoxicology*, Dublin, December 1992.

Williams, D.C., Fonnum, F., Regan, C.M., Griffiths, R., Meredith, C. & Rodríguez-Farré, E. In the field of *in vitro* screening in neurotoxicology and pharmacology using neural primary culture and cell lines *ELWW, Commission of the European Communities, DG XII* (1991)

Williams, D.C., Meredith, C., Schousboe, A., Griffiths, R & Rodríguez-Farré, E. Development of *in vitro* neural systems for the identification of agents with toxicological and pharmacological potential. *BRIDGE Progress Report* (1992)

Williams, D.C., Meredith, C., Schousboe, A., Griffiths, R & Rodríguez-Farré, E. Development of *in vitro* neural systems for the identification of agents with toxicological and pharmacological potential. *BRIDGE Progress Report* (1993)
Williams, D.C., Griffiths, R., Meredith, C., Rodríguez-Farré, E. & Schousboe, A. (eds.) *Proceedings of the CEC Biotechnology BRIDGE Programme* In vitro *Pharmacotoxicology Sectoral Meeting, Dublin, 1992* (1993)
Zisterer, D.M., Moynagh, P.N., Suñol, C., Rodríguez-Farré, E. & Williams, D.C. PK 11195 inhibits the convulsant activity of lindane by delaying its absorption into the blood. *Eur. J. Pharmacol.* (submitted).

Individual publications

Abraham, J.H., Hansen, G.H., Seiler, N. & Schousboe, A. Depletion of polyamines prevents the neurotrophic activity of the GABA-agonist THIP in cultured rat cerebellar granule cells. *Neurochem. Res.* **18**, 153-158 (1993).
Abraham, J.H., Seiler, N. & Schousboe, A. Induction of low affinity $GABA_A$-receptors by the GABA-agonist THIP (4,5,6,7-tetrahydroisoxazolo[5,4-c]pyridin-3-ol) in cultured rat cerebellar granule cells is prevented by inhibition of polyamine biosynthesis. *J. Neurosci. Res.* (in press) 1994.
Cristòfol, R. & Rodríguez-Farré, E. Modulation of noradrenaline release from hippocampal slices by hexachlorocyclohexane isomers. — Effects of GABAergic compounds. *Brain Res.* **606**, 237-243 (1993).
Cristòfol, R. & Rodríguez-Farré, E. Hippocampal noradrenaline release is modulated by γ- and δ-HCH hexachlorocyclohexane isomers: which mechanisms are involved? *Eur. J. Pharmacol.* **252**, 305-312, 1994.
Cristòfol, R. & Rodríguez-Farré, E. Differential presynaptic effects of hexachlorocyclo-hexane isomers on noradrenaline release in cerebral cortex. *Life Sciences* **49**, 1111-1119, 1991.
Cristòfol, R., Rodríguez-Farré, E. & Sanfeliu, C. Effects of γ and δ hexachlorocyclohexane isomers on inositol phosphate formation in cerebral cortex and hippocampal slices from developing and adult rat. *Neurotoxicology* **14**, 451-458 (1993).
Dunlop, J., Grieve, A. & Griffiths, R. L-*trans*-Pyrrolidine-2,4-dicarboxylic acid has properties consistent with that of a competitive substrate for the plasma membrane glutamate transporter. *Biochem. Soc. Trans.* **21**, 111S (1993).
Elster, L., Saederup, E., Schousboe, A. & Squires, R.F. Characterization and function of binding sites for ω-conotoxin in cerebellar granule neurons: Lack of interaction with $GABA_A$-receptors. *J. Neurosci. Res.* (in press) 1994.
Frandsen, Aa. & Schousboe, A. Excitatory amino acid mediated cytotoxicity and calcium homeostasis in cultured neurons. *J. Neurochem.* **60**, 1202-1211 (1993).
Frandsen, Aa. & Schousboe, A. Mobilization of dantrolene-sensitive intracellular calcium pools is involved in the cytotoxicity induced by quisqualate and N-methyl-D-aspartate but not by 2-amino-3-(3-hydroxy-5-methylisoxazol-4-yl)propionate and kainate in cultured cerebral cortical neurons. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2590-2594 (1992).
Gorman, A. & Griffiths, R. Stimulation of inositol phospholipid hydrolysis by sulphur-containing excitatory amino acids in primary cultures of cerebellar granule cells. *J. Neurochem.* **61** (Suppl.), S18D (1993).
Gorman, A. & Griffiths, R. Sulphur-containing excitatory amino acid stimulated inositol phosphate formation in primary cultures of cerebellar granule cells is mediated predominantly by N-methyl-D-aspartate receptors. *Neuroscience* **59**, 299-308, 1994.
Grieve, A. & Griffiths, R. Simultaneous measurement by HPLC of the excitatory amino acid transmitter candidates homocysteate and homocysteine sulphinate supports a predominant astrocytic localisation. *Neurosci. Lett.* **145**, 1-5 (1992).
Grieve, A., Allen, S. & Griffiths, R. β-p-Chlorophenylglutamate is not a selective blocker of L-homocysteate uptake. *Neurochem. Int.* **21(S)**, B25, (1992)
Grieve, A., Butcher, S.P. & Griffiths, R. Synaptosomal plasma membrane transport of excitatory sulphur amino acid transmitter candidates: kinetic characterisation and analysis of carrier specificity. *J. Neurosci. Res.* **32**, 60-68 (1992).
Griffiths, R. Sulphur-containing excitatory amino acids: metabolism and function, in *Drug Research Related to the Neuroactive Amino Acids*, Alfred Benzon Symposium 32, (Schousboe, A., Diemer, N.H.& Kofod, H., eds.) pp. 64-81. Munksgaard, Copenhagen, 1992.

Griffiths, R. The biochemistry and pharmacology of excitatory sulphur-containing amino acids, in *Excitatory Amino Acids; from metabolism to medicine* (Griffiths, R. & Roberts, P.J., organisers), *Biochem. Soc. Trans.* **21**, 66-72 (1993).
Griffiths, R., Butcher, S.P. & Olverman, H.J. Sulphur-containing excitatory amino acids, in *Excitatory Amino Acid Receptors: design of agonists and antagonists* (Krogsgaard-Larsen, P. & Hansen, J.J., eds.) pp. 162-182. Ellis-Horwood Ltd., UK, 1992.
Griffiths, R., Grieve, A., Allen, S. & Olverman, H.J. Neuronal and glial membrane carrier-mediated uptake of L-homocysteate is not selectively blocked by b-p-chlorophenylglutamate. *Neurosci. Lett.* **147**, 175-178 (1992)
Griffiths, R., Grieve, a. & Allen, S. Kinetic studies in *in vitro* systems reveal that β-p-chlorophenylglutamate is not a selective blocker of L-homocysteate uptake. *J. Physiol.* **459**, 57P (1993)
Griffiths, R., Dunlop, J., Gorman, A., Senior, J. & Grieve, A. L-*trans*-Pyrrolidine-2,4-dicarboxylate and *cis*-1-aminocyclobutane-1,3-dicarboxylate behave as transportable, competitive inhibitors of the high-affinity glutamate transporters. *Biochem. Pharmacol.* **47**, 267-274, 1994.
Hansen, G.H., Belhage, B. & Schousboe, A. First direct electron microscopic visualizaiton of a tight spatial coupling between $GABA_A$-reeptors and voltage sensitive calcium channels. *Neurosci. Lett.* **137**, 14-18 (1992).
Kardos, J., Elster, L., Damgaard, I., Krogsgaard-Larsen, P. & Schousboe, A. Role of $GABA_B$ receptors in intracellular Ca^{2+} homeostasis and possible interaction between $GABA_A$ and $GABA_B$ receptors in regulation of transmitter release in cerebellar granule neurons. *J. Neurosci. Res.* (in press) 1994.
Meredith, C. *In vitro* neurotoxicology research at BIBRA. *BIBRA Bull.* **33**, 33-37, 1994
Moynagh, P.N. & Williams, D.C. (1992) Stabilization of the Peripheral-type Benzodiazepine Acceptor by Specific Phospholipids. *Biochem. Pharmacol.* **43**, 1939-1945,
Moynagh, P.N., O'Neill, L.A.J. & Williams, D.C. (1992) Modulation of the Peripheral-Type Benzodiazepine Acceptor by IL-1 and PMA. *Neurochem. Int.* **21**, C63.
Moynagh, P.N., Williams, D.C. & O'Neill, L.A.J. (1992) Activation of NFkB by IL-1 in C6 glioma cells. *Neurochem. Int.* **21** C62.
Moynagh, P.N. & Williams, D.C. Molecular studies on the peripheral-type benzodiazepine acceptor. *Biochem. Soc. Trans.* **20**, 267S (1992)
Moynagh, P.N., Williams, D.C. & O'Neill, L.A.J. Interleukin-1 activates transcription factor NFkB in glial cells. *Biochem. J.* **294**, 343-347 (1993)
Moynagh, P.N., O'Neill, L.A.J. & Williams, D.C. Interleukin-1 and phorbol myristate acetate modulate the peripheral-type benzodiazepine receptor in lymphocytes and glial cells. *Biochem. Pharmacol.* **46**, 821-827.
Moynagh, P.N., Williams, D.C., Gearing, A.J.H. & O'Neill, D.C. Induction of VCAM and ICAM-1 in human glioma cells by IL-1. *Lymphokine and Cytokine Research* **12**, 224 (1993)
Moynagh, P.N., Williams, D.C. & O'Neill, L.A.J. Activation of NFκB and induction of VCAM-1 and I-CAM-1 expression in human glial cells by IL-1: Modulation by antioxidants *J. Immunol.* (in press) 1994.
Nissen, J., Schousboe, A., Halkier, T. & Schousboe, A. Purification and characterization of an astrocyte GABA-carrier inducing protein (GABA-CIP) released from cerebellar granule cells in culture. *Glia* **6**, 236-243 (1992).
Pomés, A., Rodríguez-Farré, E. & Suñol, C. Interaction of polychlorocycloalkane insecticides with the $GABA_A$ receptor in primary neuronal cultures. *Neurotoxicology* **12(4)**, 799 (1991)
Pomés, A., Rodríguez-Farré, E. & Suñol, C. Inhibition of *t*-[^{35}S]butylbicyclophosphorothionate binding by convulsant agents in primary cultures of cerebellar neurons. *Dev. Brain Res.* **73**, 85-90 (1993).
Pomés, A., Rodríguez-Farré, E. & Suñol, C. Effects of organochlorine pesticides on ^{36}Cl- flux in primary neuronal cultures. *Neurotoxicology* **15(1)**, 001-005 (in press) 1994.
Pomés, A., Rodríguez-Farré, E. & Suñol, C. Inhibition of GABA-induced ^{36}Cl- uptake in primary cultures of cortical neurons by convulsant organochlorine pesticides. 16th Annual Meeting of the European Neuroscience Association. Madrid, Spain, 1993.
Pomés, A., Rodríguez-Farré, E. & Suñol, C. Disruption of GABA-dependent chloride flux by cyclodienes and hexachlorocyclohexanes in primary cultures of cortical neurons. *J. Pharmacol. Exp. Ther.* (submitted).

Pomés, A., Rodríguez-Farré, E. & Suñol, C. GABA-dependent Cl- flux in primary cultures of cortical neurons. Effects of convulsant/depressant agents. *International Meeting on Ion Channel Pharmacology*. Alicante, Spain, April 1994.
Rodríguez-Farré, E., Roberfroid, M. & Fracchia, G.N. Research and development of *in vitro* pharmacology — A European perspective *ATLA*, **21(2)**, 285-293 (1993)
Sanfeliu, C., Cristòfol, R.M. & Rodríguez-Farré, E. Effects of hexachlorocyclohexane isomers on inositol phosphate formation in cerebral cortex and hippocampus slices from developing and adult rat. *Neurotoxicology* **13(4)**, 877, 1992.
Sanfeliu, C., Cristòfol, R.M., Rosa, R. & Rodríguez-Farré, E. γ- and δ-hexachlorocyclohexane isomers stimulate inositol phosphate formation in cerebellar granule cells. *Neurotoxicology* **14**, 555, 1993.
Sanfeliu, C., Cristòfol, R.M., Rosa, R. & Rodríguez-Farré, E. γ- and δ-hexachlorocyclohexane isomers stimulate inositol phosphate formation in cerebellar granule cells. 4th Meeting of the International Neurotoxicology Association. Elsinore, Denmark, 1993.
Suñol, C., Pomés, A. & Rodríguez-Farré, E. Primary neuronal cultures to study the $GABA_A$ receptor as a target for neurotoxic agents: different effects of stimulant/depressant compounds. In: Perspectives in Pharmacotoxicology and Pharmacovigilance. (G. Fracchia, ed.) *IOS Press*, Amsterdam 1994.
Wahl, P., Elster, L. & Schousboe, A. Identification and function of glycine receptors in cultured cerebellar granule cells. *J. Neurochem.* (in press) 1994.
Westergaard, N., Larsson, O.M., Jensen, B. & Schousboe, A. Synthesis and release of GABA in cerebral cortical neurons co-cultured with astrocytes from cerebral cortex or cerebellum. *Neurochem. Int.* **20**, 567-575 (1992).
Westergaard, N., Varming, T., Peng, L., Sonnewald, U., Hertz, L. & Schousboe, A. Uptake, release, and metabolism of alanine in neurons and astrocytes in primary cultures. *J. Neurosci. Res.* **35**, 540-545 (1993).
White, H.S., Hunt, J., Wolf, H.H., Swinyard, E.A., Falch, E., Krogsgaard-Larsen, P. & Schousboe, A. Anticonvulsant activity of the g-aminobutyric acid uptake inhibitor N-4,4-diphenyl-3-butenyl-4,5,6,7-tetrahydroisoxazolo[4,5-c]pyridin-3-ol. *Eur. J. Pharmacol.* in press (1993).
Williams, D.C. Biotechnology Programme: Sectoral Meeting on *In vitro* Evaluation of the Toxicity and Pharmacological Activity of Molecules. *CEC Biotech BRIDGE Programme Sectoral Meeting on* in vitro *evaluation of the toxicity and pharmacological activity of molecules. ATLA* (1993)
Woods, M.J., Moynagh, P.N. & Williams, D.C. Rat adrenal and kidney peripheral-type benzodiazepine receptors are primarily located in contact sites *Eur. J. Biochem.* submitted (1994)
Zisterer, D.M., Gorman, A.M.C., Williams, D.C. & Murphy, M.. (1992) The Effects of the Peripheral-type Benzodiazepine Receptor Ligands, Ro5-4864 and PK 11195 on Mitochondrial Respiration. *Methods Findings Exp. Clin. Pharmacol.* **14**, 85-90.

Definition and scientific validation of an *in vitro* test for the screening of tumor promoters and protective chemicals, based on gap junctional intercellular communication assays in human and animal cells (BIOT CT-910261)

COORDINATOR:
E. HONIKMAN-LEBAN, L'Oréal, Aulnay-sous-Bois, F
PARTICIPANTS:
M. BALLS, FRAME, Nottingham, GB
E. HONIKMAN-LEBAN, L'Oréal, Aulnay-sous-Bois, F
P. MARTEL, INRA, Jouy-en-Josas, F
G. MAZZOLENI, University of Brescia, Brescia, I
M. MESNIL and V. KRUTOVSKIKH, IARC, Lyon, F
G. RAGNOTTI, University of Milan, Milan, I
L. WARNGARD, Karolinska Institute, Stockholm, S
K. WILLECKE, University of Bonn, Bonn, D

METHODOLOGY

GJIC can be established in cell culture monolayers at a confluent stage, and measured by the *dye transfer assay*. The technique selected in our project is the microinjection of vital, fluorescent dye (Lucifer Yellow): the dye, loaded in a microcapillary, is automatically monitored by micromanipulator and microinjector connected to an inverted fluorescent microscope. The microinjection is performed in a single cell and the diffusion of the dye through gap junctions is quantified by the fluorescence spreading to the adjacent cells. When GJIC is inhibited, no diffusion occurs and only the microinjected cell is fluorescent.

RESULTS AND DISCUSSION

The first part of our programme focused on the *standardization of the dye transfer assay*, using IAR20 cells, by testing five known chemicals at three concentrations and four treatment times, in order to set up the best protocol between the participants.
The second period focused on:
the development of other cells types, using molecular tools to identify specific GJIC proteins (cx), the study of cx gene expression regarding the effects of tumor promoter treatments (HeLa cx-transfectants);
cytotoxicity evaluation for compounds being tested with dye transfer assay (MTT, neutral red uptake, cell growth, cloning efficiency).
The last part of our programme was devoted to the **Pre-validation Blind Trial Experiments**: the interlaboratory pre-validation study was conducted, as designed by **L'Oréal** in consultation with **FRAME** and involving the independent purchase, coding and distribution of test materials by BIBRA Toxicology International (Carshalton, UK) and the independent receipt and preliminary analysis of the results by ECVAM (Ispra, Italy). The test chemicals were chosen by FRAME with the other participating BRIDGE project partners. Each of the five participating laboratories (L'Oréal, INRA, IARC, Karolinska I., Brescia U.) received 16 test materials, with a unique code being used for each laboratory. The blind trial was performed following the protocol proposed in Stockholm and drawn up by coordinator, Dr. Esther Honikman-Leban:

— IAR 20 cells were used at passage numbers from 26 to 30. The plating density was 5×10^5 cells/plate (∅ 60 mm) and cells were maintained in 3 ml W-MEM + 10% FCS.
— All the compounds to be tested were dissolved in the solvent indicated (by BIBRA) at the concentration of 10 µg/ml.
— The solvent was tested up to 1% v/v and it wad confirmed that it had no effect on IAR 20 cells GJIC.
— Cells were treated for 1, 4, 24 hr. and then microinjected during a 6 min. period (almost 30 microinjections) and fixed with 4% fresh prepared formaldehyde 5 min after the last microinjection. Plates were then refrigerated. Two independant experiments for each experimental point (one compound at one concentration at one time) were performed, each experiment including two plates from which 15 'successful' microinjections were scored.

The chemicals chosen were as follows: 12-O-tetradecanoylphorbol-13-acetate (TPA); -phorbol-12,13-didecanoate (4α-PDD); phorbol-12, 13-dibutyrate (PDBu); stearic acid; palmitic acid; palmitelaidic acid; butylated hydroxyanisole (BHA); saccharin; retinoic acid; phenobarbital; diethylstilbestrol; benzoyl peroxide; DDT; warfarin, D-(ethylhexyly)phthalate (DEPH) and 4-O-methylphorbol,12-myristate,13-acetate (MPMA). The results of this pre-validation blind trial are still being analysed, but it can already be said that the outcome was very encouraging and that **a formal, full validation study would be worthwile.** For example, 5 of the chemicals tested clearly did not affect gap junctional intercellular communication (GJIC) when tested in any of the five laboratories. Of the remaining 11 test materials, 8 were shown to inhibit GJIC by three or more laboratories, while the remaining 3 were shown to be inhibitory by one or two laboratories. Many useful lessons were learned, about the test protocol, about the selection and supply of chemicals, and about the analysis of the test data!

L'OREAL Laboratory participated in the standardization of the dye transfer assay with IAR20 cells, and in the work on the cytotoxicity evaluation for compounds being tested with dye transfer assay using neutral red, trypan blue with IAR20 cells and cloning efficiency with V79 cells. In collaboration with **INRA** and **IARC** laboratories, experiments were completed to test the modulating (anti-promoting) effects of flavonoids, vitamins A and E and related compounds, phenolic anti-oxydants (BHA, BHT) using metabolic cooperation assay with V79 cells (Chinese hamster lung fibroblasts), the dye transfer techniques with IAR20 cells and human keratinocytes: scrape-loading and microinjection as well as FRAP (photobleaching). The results of these studies are described in collaborative publications. These *in vitro* techniques are complementary for testing compounds directly or in combination with TPA (potent tumor promoter, used as reference in GJIC). The V79 system is able to show synergistic or antagonistic effects on GJIC inhibition, but unable to show the enhancement of GJIC. The microinjection assay is able to show this and has the advantage of permitting to work with different, more suitable cell types. This system showed a clear improvement with respect to metabolic cooperation with V79 cells. During this project, we attempted to quantify automatically the intercellular transfer of fluorescence using Image Analysing techniques. A major part of our contribution covered coordination work and the organization of the Blind Trial experiments (reported above).

INRA Laboratory, together with the other participants, contributed to the standardization of the dye transfer assay, and to the evaluation of *in vitro* cytotoxicity assays (cell density determination and neutral red uptake) to be applied to inhibitors of intercellular communication.

In addition, the ability of various compounds to modulate gap junctional intercellular communication (GJIC) of liver cell systems was studied at the functional and molecular level. TPA, DDT, phenobarbital (PB) were used as reference tumor promoters. The other compounds chosen were food additives (BHA, BHT), vitamins (retinoic acid, retinol acetate, α-tocopherol, tocopherol succinate) and natural flavonoids (12 compounds taken in the different flavonoid classes). With IAR 20 cells (clone 203), retinol acetate and tocopherol succinate inhibited GJIC, whereas retinoic acid was able to enhance or inhibit it, depending on the concentration, and α-tocopherol was neutral. The mechanism of the stimulation of GJIC by retinoic acid was studied: it increased the expression of connexin 43 at the post-transcriptional level. With another rat liver epithelial cell line isolated in the laboratory, named REL cells, the activity of the different flavonoids on GJIC was investigated. Four flavonoids, including quercetin and kaempferol, were inactive. A slight increase of GJIC was observed with six compounds, and a marked one with apigenin and tangeretin. Apigenin and tangeretin enhanced connexin 43 protein levels and were also able to antagonize the inhibition of GJIC by TPA. All the tumor promoters tested (TPA, DDT, PB, BHA, BHT) inhibited GJIC of both IAR 203 cells and REL cells. Other cell systems were developed and applied to the study of the regulation of connexin 43 and 32 expression. REL cells were transfected with cDNAs encoding c-fos and c-jun oncoproteins. Proliferation, transformation and homologous GJIC of transfectants were studied in comparison with parental cells. These models provided evidence of the up-regulation of connexin 43 by c-fos overexpression and its down-regulation by c-jun overexpression. Their GJIC were affected differently by tumor promoters. They are good tools for studying the involvement of some nuclear oncogenes in the mechanism of inhibition of GJIC by tumor promoters.

IARC Laboratory emphasized its effort in the confirmation of cx26 as a tumour suppressor gene since our previous data suggested that connexin 26 transfected into non-communicating HeLa cells is able to reduce their growth rate. In collaboration with **Bonn laboratory**, we could obtained other clones of HeLa cells transfected with the cx26 cDNA and compare their growth ability with HeLa cells transfected with other connexin types (cx40 and 43). Two clones out of three expressing the highest levels of cx26 exhibited a marked inhibition of growth rate in the *in vitro* conditions contrary to the other clones expressing other connexins or less cx26. Not only cell growth, but also cell density, were decreased in these two clones. They also exhibited a very decreased ability to grow in suspension in soft agar, which is a capacity of transformed cells. Moreover, their tumorigenicity, which was completely inhibited in nude mice, confirmed our preliminary results suggesting that cx26 is a putative tumor suppressor. We first hypothesized that the lack of tumorigenicity of the cx26 transfectants could be due to the establishment of communication with the surrounding normal cells of the nude mouse dermis. However, the *in vitro* search for such communication between the non-tumorigenic cx26 transfectants and primary cultures of keratinocytes and skin fibroblasts was negative when the dye transfer technique was used. Thus, cx26 seems to regulate intrinsinc the growth of the HeLa cells both *in vitro* and *in vivo* situations.

The interesting aspect of the study is that although transfectants expressing other types of connexins (cx40 and 43) exhibited similar dye-transfer ability to that in the cx26 transfectants, they do not exhibit any decreased growth. From these data the induced cell-cell communication capacity does not seem to be a prerequesite for the down-regulation of the cell growth. However, it is still possible that the dye-transfer technique is not sufficiently sensitive to detect subtle differences of communication which could be mediated by the different connexins and which could

have significant consequences on the cell behaviour. The specific effect of cx26 expression may be related to the tissue origin of the HeLa cells, which were isolated from an adenocarcinoma of the cervix. We are preparing tools for preventing the synthesis of cx26 in order to see whether transfectants recover then a growth ability similar to that of non-transfected HeLa cells. The transfectants were used to see the effect of different tumour promoters on the GJIC mediated by the different transfected connexins. We conducted this work in collaboration with **Brescia laboratory**. Our part was to demonstrate the modulation of the mRNA expression due to the treatments in such transfectants. Finally, in an attempt to compare *in vitro* and *in vivo* situations, we studied the effects of potent rat liver tumor-promoting agents on the GJIC of rat liver by the *in situ* dye transfer technique, which demonstrated that the induced inhibition of GJIC is mostly the consequence of a delocalisation of connexins which are expressed in the hepatocytes (cx32 and 26), but not a lack of expression. All these results taken together emphasize the important role that GJIC may play in carcinogenesis by showing the sensitivity of connexins to tumour promoting agents both *in vitro* and *in vivo* conditions and their possible tumour suppressor effects.

KAROLINSKA Laboratory has been working on three groups of substances, the phorbol ester TPA, polychlorinated biphenyls (PCBs) and, finally, the insecticide endosulfan. These chemicals have been used since they are structurally different and induce different biochemical effects in biological experimental systems. Endosulfan and TPA were studied in IAR 20 cells, which are used within the project as a test strain. The results clearly demonstrate that both compounds inhibited cell-cell communication within 5 minutes. The communication was restored after 4 hours of TPA treatment, whereas in endosulfan-treated cells the communication was completely downregulated during the whole time period (24h). Using an immunoblotting technique combined with immunostaining, it was shown that TPA caused a marked hyperphosphorylation, whereas endosulfan induced hypophosphorylation. These results suggest that these chemicals inhibit the GJIC by different mechanisms. The expression of cx32 and cx26 in rat livers treated with 3,4,5,3',4', -PCB in a tumour promotion assay has been investigated with immunoblotting as well as with the immunostaining of liver sections. The PCB-congener induced a potent reduction of the amount of bound fraction cx32 and of 26. The immunostaining showed a nearly 90% reduction of apparent connexins in the membrane. The amounts of cx32- and cx26-mRNA were unaffected. This work was done in collaboration with **IARC** and **Bonn laboratory**. In collaboration with the **Bonn laboratory** we are investigating the effect of certain tumour promoters on GIJC using HeLa cx43, 32 and 26 transfectants, as well as primary hepatocytes, in order to study connexin-specific effects. We have started with two groups of chemicals DDT-analogs and pyrethroids, both groups are used as pesticides. These chemicals have been shown to decrease gap junction expression *in vivo* and *in vitro*. The results from these studies shows that there are differences in regulation and in sensitivity to chemically-induced inhibition of gap junctional communication among the systems used.

BRESCIA — MILAN Laboratory evaluated the effects of tumor-promoting (TPA, PB), non-promoting (4-α-PDD) and anti-promoting (RA, dbcAMP) agents on human HeLa cells (not expressing any known GJ protein and constitutively deficient in GJIC) transfected with cDNAs coding for different cxs: cx43, cx26, cx40 and cx32 (**Bonn** collaboration), in order to investigate the implication of GJIC in the control of cell growth (**IARC** collaboration) and to compare the properties of homotypic junctional channels. The GJIC capacity of the transfected clones was ≃ 40 (HeLa-cx43 and HeLa-cx32), ≃ 20 (HeLa-cx26) and ≃ 15 (HeLa-

cx40) dye-coupled cells/injection. The effects on cell viability (proliferation) or junctional coupling (microinjection/dye-transfer assay) of 4, 24 and 48 hr exposure to different concentrations of the tested compounds were studied. The results obtained indicate that the cx-transfectants have a different susceptibility to the cytotoxic effect exerted by the chemicals, HeLa-cx26 being generally more susceptible than the other HeLa-transfectants or the parental HeLa cells. According to the cx-transfectant analysed, the GIJC response to the compounds was also somehow different. TPA was found to induce a significant reduction of the dye-coupling capacity of the transfectant cells capacity which, 48 hr after the treatment, returned to control values in the case of cx43 and cx40-transfectants. HeLa-cx26 GJIC was maximally inhibited during all the 48 hr period of treatment, while TPA reduces HeLa-cx32 GJIC to 50% of control values only after 48 hrs of treatment. The transfectant junctional coupling was only slightly modified by the other chemical tested, the modifications being time of exposure-dependent and cx-dependent, but generally, not comparable in extent with the GJIC down-modulation caused by TPA.
Our results seem to confirm that the cx diversity may be responsible for the different gating properties of gap-junctional channels and that they may have separate functions, can be independently regulated and respond differently to signals. Immunoblot analysis and light microscopy studies are in progress, in order to clarify the role of post-translational modifications in the observed TPA-induced inhibition of the transfectant junctional coupling we have observed.

BONN Laboratory has completed the following studies involving antibodies to connexins(cx) 40 and -43. They have shown that cx40 is specifically localized in conducting myocardiocytes of rat heart and to a much lesser extent in working myocardium -in contrast to cx43 which is mainly found in the latter part of rat heart. In hypertensive rat heart, the amount of cx40 is increased threefold whereas cx43 decreases under the same conditions. The cx40 antibodies have been used for characterization of this protein in transfected HeLa cells. Furthermore, the cx45 protein has been shown to be expressed in glomeruli and distal tubules of kidney in 4 day old mice.
Specific connexin antibodies were used to study differential expression of the gap junction proteins cx45, cx43, cx40, cx31 and cx26 in mouse skin. In a collaborative study with the **IARC laboratory**, it was shown that the HeLa cx26 transfectants exhibit reduced tumorigenicity in nude mice when compared with parental HeLa cells or other HeLa connexin transfectants. Furthermore, we have shown that out of seven different HeLa connexin transfectants, only few can form functional heterotypic channels, suggesting that incompatible connexins may contribute to communication barriers in early mammalian development (Elfgang *et al.*, in preparation, 1994). The connexin gene probes and antibodies characterized in this laboratory have been sent to four laboratories (**Brescia U., IARC, INRA, Karolinska I.**) collaborating in the BRIDGE programme for the characterization of the effects of tumor promoters at the level of connexin mRNAs and proteins.

COOPERATIVE ACTIVITIES

Meetings and interactions with other BRIDGE groups:
- **BAMMIF**, London, GB, 21-24/04/92: First European Workshop on Biotechnology Applications of Microinjection, Microscopic Imaging and Fluorescence.
- **EC-BRIDGE sectoral meeting**, Dublin, IR, 8-10/12/92: In Vitro Evaluation of Toxicity and Pharmacological Activity of Molecules (Part I).
- **International meeting on Gap Junctions**, Hiroshima, JN, 24-28/08/93: Presentation of GJIC-BRIDGE programme and results by some of the project leaders.

— **ELWW joint meeting** in the field on *In Vitro Integrated Approach to Skin Pharmaco-toxicology* and in the field of *Gap Junctional Intercellular Communication and Screening of Tumor Promoters, Montpellier*, F, 25-26/11/93: Importance of Cellular Interactions in the Skin Pharmaco-Toxicology.
— **EC-BRIDGE sectoral meeting,** Innsbruck/Igls, AU, 5-8/04/94: *In Vitro* Evaluation of Toxicity and Pharmacological Activity of Molecules (Part II).

Exchanges and collaborations:
— **Bonn University — Brescia U. — IARC — INRA — Karolinska Institute:** exchanges of antibodies specific to different connexins and protocols.
— **Bonn University — IARC** (1992-1993): study of growth and tumorigenicity of HeLa cells transfected by cx genes. A collaborative paper is submitted to Cancer Research.
— **Brescia-Milan Universities — IARC** (1994): study of the sensibility of the cx gene (mRNA) transfected HeLa cells towards tumor promotion treatments (PB, TPA, etc,...).
— **Brescia-Milan Universities — Bonn University** (1993-1994): exchanges of cx gene transfected HeLa cells.
— **IARC — Karolinska Institute** (1991-1992): exchange of techniques and collaborative work on the effects of PCBs on cx function and expression in liver of treated rats. A collaborative paper is in preparation.
— **INRA — L'Oréal** (1992-1994): Comparison of the effects of different compounds (Vitamines A and E) on intercellular communication, with two functional assays: dye transfer and metabolic cooperation assay (papers are in preparation).
— **INRA — IARC — L'Oréal** (1992-1994): Comparison of the effects of various flavonoids on intercellular communication, *in vivo* and *in vitro*. A collaborative paper is submitted to Nutrition and Cancer.
— **L'Oréal — Brescia U. — FRAME — IARC — INRA — Karolinska Institute:** coordination of:
  — meetings and joint experiments (1991-1994);
  — cytotoxicity (1993) and
  — blind trial test: protocol, cells and compounds exchanges. (1994).

PUBLICATIONS

Joint publications

Meetings (abstracts and papers)
London (April 92): BAMMIF organized by P. Bach (BIOT-CT91-0266); abstracts of the workshop have been published, In: Biotechnology Applications of Microinjection, Microscopic Imaging and Fluorescence, P.H. Barch & al (Eds), *Plenum Press*, NY, 1993.
Dublin (Dec. 92): EC-BRIDGE sectoral meeting organized by Cl. Williams (BIOT-CT90-0183) with the BRIDGE groups of the pre-normative research area; highlights of the meeting will be published in *ATLA* (1993-94).
Hiroshima (August 93): International meeting on gap junctions; BRIDGE results will be published by project leaders, In: Gap Junctions, Y. Kanno (ed.), *Elsevier Science*, BV. Amsterdam, AM, 1994. In press.
Montpellier (Nov. 93): ELWW meeting organized by A. Barbier, B. Coulomb and E. Honikman-Leban (BIOT-CT90-0186 , BIOT-CT90-0193 and BIOT-CT91-0261), *abstracts booklet* available c/o A. Barbier (Sanofi).
Innsbruck/Igls (April 94): EC BRIDGE sectoral meeting, organized by W. Pfaller (BIOT-CT91-0266) with the BRIDGE groups of the pre-normative research area; *highlights* of the meeting will be published in *Journal of Cellular Physiology and Biochemistry*, 1994.

BRIDGE project publications

BAGER Y., KENNE K., KRUTOVSKIKH V., MESNIL M. , TRAUB O., YAMASAKI H. and WARNGARD L. (1994) Alteration in expression of gap junction proteins in rat liver after treatment with 3,4,5,3',4'-Petachlorobiphenyl (in preparation).

BEX V., HONIKMAN-LEBAN E., MAZET F., FLECHON B., MERCIER T., SHAHIN M. M. and MARTEL P., *In vitro* modulation of gap junctional intercellular communication by retinoic acid: functional, molecular and structural studies, VI International Congress of Toxicology, Rome, June 28-July 3, 1992.

CHAUMONTET C., SUSCHETET M., HONIKMAN-LEBAN E., KRUTOVSKIKH V., LEBON A.M., BERGES R., SHAHIN M.M., YAMASAKI H. and MARTEL P. (1994) Lack of liver tromotingts of flavonoids. *In vivo* and *in vitro* studies of their action on gap junctional intercellular communication, submitted for publication to Nutrition and Cancer.

FRANSSON-STEEN R., ELFGANG C., WILLECKE K. and WARNGARD L. (1994) Inhibitory properties on gap junctional communication by DDT and its analogues in cultured cells expressing different connexins (Manuscript).

KRUTOVSKIKH V., MESNIL M., MAZZOLENI G. and YAMASAKI H. (1994) Inhibition of rat liver gap junction intercellular communication by tumor-promoting agents *in vivo*. Association with aberrant localization of connexin proteins, Lab. Investigations (submitted).

KRUTOVSKIKH V., MAZZOLENI G., MIRONOV N., OMORI Y., AGUELON AM., MESNIL M., BERGER F., PARTENSKY C. and YAMASAKI H. (1994) Altered homologous and heterologous gap junctional intercellular communication in primary human liver tumors associated with aberrant protein localization but not gene mutation of connexin 32, *Int. J. Cancer*, 56, 87-94.

MAZZOLENI G., CAMPLANI A., TELO' P., MESNIL M., WILLEKE K., TANGANELLI S., POZZI A., and RAGNOTTI G. (1994) TPA differentially affects the permeability of homotopic gap-junctional channels: an *in vitro* study on HeLa cells transfected with various connexin genes (in preparation).

MERCIER T., HONIKMAN-LEBAN E., CHAUMONTET C., MARTEL P. and SHAHIN M.M. (1993) Studies on the modulating effects of retinoic acid and retinol acetate using dye transfer and metabolic cooperation assays, *Fund. Appl. Toxicol.*, **21**, 270-276.

MESNIL M., KRUTOVSKIKH V., PICCOLI V., ELFGANG C., TRAUB O., WILLECKE K. and YAMASAKI H. (1994) Negative growth control of HeLa cells by connexin genes: connexin species specificity (submitted).

MESNIL M., KRUTOVSKIKH V., PICCOLI C., ELFGANG C., WILLECKE K. and YAMASAKI H. (1994) Negative growth control of HeLa cells by connexin genes: connexin-species specificity, Cancer Res. (submitted).

ALL PROJECT LEADERS (1994-95) Special Issue on 'GJIC and Tumor Promotion' (in preparation).

Individual publications

BASTIDE B., NEYSES L., GANTEN D., PAUL M., WILLECKE K. and TRAUB O. (1993) Gap junction protein connexin 40 is preferentially expressed in vascular endothelium and conductive bundles of rat myocardium and is increased under hypertensive conditions, *Circulation Research*, **73**, 1138-1149 .

BUTTERWECK[1] A., GERGS U., ELFGANG C., WILLECKE K. and TRAUB O. (1994) Immunochemical characterization of the gap junction protein connexin 45 in mouse kidney and transfected human cells (submitted).

BUTTERWECK[2] A., ELFGANG C., WILLECKE K. and TRAUB O. (1994) Differential expression of the gap junction proteins Cx45, Cx43, Cx40, Cx31 and Cx26 in mouse skin (submitted).

CHAUMONTET C., BEX V., GAILLARD-SANCHEZ I., SEILLAN-HEBERDEN C., SUSCHETET M. and MARTEL P. (1994) Apigenin and tangeretin enhance gap junctional intercellular communication in rat liver epithelial cells, *Carcinogenesis* (in press).

COOPER A., FENTEM J. and BALLS M. (1994) Inhibition of intercellular communication *in vitro* assay for detecting potential tumor promoters, *Toxicology in Vitro*, **8**, 535-537.

ELFGANG C., LICHTENBERG-FRATE H., BUTTERWERCK A., ECKERT R., TRAUB O, KLEIN R., HÜLSER D. and WILLECKE K. (1994) Expression of murine connexin DNAs in human HeLa cells shows differences in permeabilities of different tracers and selectivity in the formation of functional gap juncton channels (submitted).

FITZGERALD DJ., SWIERENGA SHH., MESNIL M., PICCOLI C., MARCEAU N. and YAMASAKI H. (1993) Gap junctional intercellular communication and connexin expression in normal and SV40-transformed human liver cells *in vitro, Cancer Lett.,* **71**, 157-165.
HEMMING H., FLODSTROM S. and WARNGARD L. (1994) Enhancement of altered hepatic foci in rat liver and inhibition of intercellular communication *in vitro* by the synthetic pyrethroids fenvalerate, flucythrinate and cypermethrin, *Carcinogenesis,* **14**, 2531-2535.
HONIKMAN-LEBAN E. and SHAHIN M.M. (1992) Gap junctional intercellular communication and screening of potential tumor promoters, *In Vitro Cellular and Developmental Biology,* **28**, 138A.
HONIKMAN-LEBAN E., CATROUX P., METEZEAU P., KIEFER H., BAUMAN I., COTTIN M., ROUGIER A., DOSSOU K.G. and SHAHIN M.M. (1993) Possible application of image laser cytofluorometry in cellular dye transfer and determination of free calcium, *Cell Biol.,* **76**, 39.
HONIKMAN-LEBAN E. and SHAHIN M.M., Comparison of fluorescent dye transfer in intercellular communication studies using scrape-loading and microinjection techniques, In: Biotechnology Applications of Microinjection, Microscopic Imaging and Fluorescence, P.H. Barch & al (Eds), Plenum Press, NY, 1993, p 45-52.
HONIKMAN-LEBAN E. and SHAHIN M.M., Gap junctional intercellular communication: an assay for screening of tumor promoters and 'antipromoters', in: Sectoral Meeting on *in vitro* Evaluation of the Toxicity and Pharmacological Activity of Molecules, Dublin 8-10 December 1992.
HONIKMAN-LEBAN E. and SHAHIN M.M., Gap junctional intercellular communication: a european collaborative study, In: Gap Junctions, Y. Kanno (ed.), Elsevier Science, BV Amsterdam, AM, 1994 (in press).
HONIKMAN-LEBAN E. and SHAHIN M.M., Gap junctional intercellular communication: a european collaborative study, in: Sectoral Meeting on *in vitro* Evaluation of the Toxicity and pharmacological Activity of Molecules, Innsbruck/Igls, 5-8 April 1994.
KATOH F., KLEIN JL., BIGNAMI M. and YAMASAKI H. (1993) Association of viral oncogene-induced changes in gap junctional intercellular communication and morphological transformation in BALB/c 3T3 cells, *Carcinogenesis,* **14**, 435-440.
KENNE K., FRANSSON-STEEN R., HONKASALO S. and WARNGARD L. (1994) Two inhibitors of gap junctional intercellular communication, TPA and endosulfan: Different effects on connexin 43 in the rat liver epithelial cell line, IAR20. Accepted in *Carcinogenesis.*
LAGARRIGUE S., SEILLAN-HEBERDEN C., MARTEL P. and GAILLARD-SANCHEZ I. (1992) Altered response to growth factors in rat epithelial liver cells overexpressing human c-fos protein, *FEBS Lett.,* **314** (3), 399-403.
LINNAINMAA K., PELIN K., VANHALA E., TUOMI T., PICCOLI C., FITZGERALD DJ. and YAMASAKI H. (1993) Gap junctional intercellular communication of primary and asbestos-associated malignant human mesothelial cells, *Carcinogenesis,* **14**, 1597-1602.
MAZZOLENI G., CAMPLANI A., TELO' P., TANGANELLI S. and RAGNOTTI G., The dye-transfer assay permits the evaluation of the modulation of junctional communication by chemicals: results obtained employing a highly sensitive video-recording system connected with the microinjector, In: 'Microscopic Imaging, Fluorescence and Micro-injection in Biotechnology', P.H. Bach, C.H. Reynolds, J.M. Clarck, J. Mottley e P.L. Poole eds, Plenum Press, New York, N.Y., 1993, p.37-43.
MAZZOLENI G., TELO' P., CAMPLANI A., TANGANELLI S., MONARCA S. and RAGNOTTI G. (1994) Influence of the Herbicide Linuron on the growth Rate and Gap-Junctional Intercellular Communication Capacity of Cultured F-BAE GM 7373 Endothelial Cells, In: Journal of Environmental Toxicology, Pathology and Oncology, CRC Press (in press).
MAZZOLENI G., CAMPLANI A., TELO' P., POZZI A., TANGANELLI S. and RAGNOTTI G. (1994) Is the intracellular localization of connexin 43 a possible mechanism for the TPA-induced inhibition of rat liver epithelial cells junctional coupling ?, In: Gap Junctions, Y. Kanno (ed.), Elsevier Science, BV Amsterdam AM, (in press).
MERCIER T., GAILLARD-SANCHEZ I., MARTEL P. and SEILLAN-HEBERDEN C. (1994) Constituve overexpression of c-fos protein in rat liver epithelial cells decreases TGF-α_1 synthesis and increases TGF-α_1 receptors (submitted).
MESNIL M. and YAMASAKI H. (1993) Cell-cell communication and growth control of normal and cancer cells: evidence and hypothesis, *Mol. Carcinogenesis,* **7**, 14-17.

MESNIL M., OYAMADA M., FITZGERALD DJ., JONGEN WMF., KRUTOVSKIKH V. and YAMASAKI H. (1993) Gap junctional communication alterations at various regulatory levels of connexin expression and function during animal and human carcinogenesis, In: Progess in Cell Research, vol. 3, J.E. Hall, G.A. Zamphigi and RM. Davis (Eds), Elsevier, p. 311-316.
MESNIL M., PICCOLI C. and YAMASAKI H. (1993) An improved long-term culture of rat hepatocytes to detect liver tumour-promoting agents: results with phenobarbital, *Eur. J. Pharmacol., Environ. Toxicol. Pharmacol. Section,* **248**, 59-66.
MESNIL M., PICCOLI C., KLEIN JL., MORAND I. and YAMASAKI H. (1993) Lack of correlation between the gap junctional communication capacity of human colon cancer cell lines and expression of the DCC gene, a homologue of a cell adhesion molecule (N-CAM), Jpn. *J. Cancer Res.,* **84**, 742-747.
MESNIL M. and YAMASAKI H. (1994) Le cancer, un problüme de communication? *La Recherche,* **261**, 76-78.
MESNIL M., ASAMOTO M., PICCOLI C. and YAMASAKI H. (1994) Possible molecular mechanisms of loss of homologous and heterologous gap junctional intercellular communication in rat liver epithelial cell lines, *Cell Adhesion and Communication* (submitted).
RIVEDAL E., YAMASAKI H. and SANNER T. (1994) Inhibition of gap junctional intercellular communication in Syrian hamster embryo cells by TPA retinoic acid and DDT, *Carcinogenesis,* **15**, 689-694.
TRAUB O., LICHTENBERG-FRATE H., ECKERT R., BASTIDE B., SCHEIDTMANN K.H., ELFGANG C., HUELSER D. and WILLECKE K. (1994) Immunochemical and electrophysiological characterization of murine connexins 40 and -43 in mouse tissues and transfected human cells, (submitted).
WARNGARD L., BAGER Y., HEMMING H., HONKASALO S., KENNE K., The effect of 3,4,5,3',4'-Pentachlorobiphenyl and 2,3,7,8-Tetrachlorodibezo-p-dioxin on gap junctional intercellular communication *in vitro* and *in vivo*, In: Gap Junctions, Y. Kanno (ed.), Elsevier Science, BV Amsterdam AM, 1994, (in press).
YAMASAKI H., KRUTOVSKIKH V., MESNIL M. , COLUMBANO A., TSUDA H. and ITO N. (1993) Gap junctional intercellular communication and cell proliferation during rat liver carcinogenesis, *Environ. Health Perspect,* **101** (suppl. 5), 191-198.

Development of a predictive *in vitro* test for detection of sensitizing-compounds (BIOT CT-900186)

COORDINATORS:
A. BARBIER/E. RIZOVA, Sanofi Recherche, Montpellier, F
PARTICIPANTS:
J. BOS, Departm. of Dermatology, Uva, Amsterdam, NL
J. KNOP, Department of Dermatology, Mainz, D
J.P. LEPOITTEVIN, Lab. de Dermatochimie, ULP, Strasbourg, F

BACKGROUND INFORMATION

Allergic Contact Dermatitis (ACD) belongs to Delayed-Type Hypersentivity reactions. The prevalence of contact dermatitis in the general population varies between 1.5 and 5.4% but no specific cure is currently available. Prevention of sensitization is actually the only weapon against ACD. The list of skin sensitizers is growing every day with new substances (natural or synthetic) entering the composition of household products, cosmetics or drugs.

OBJECTIVES AND PRIMARY APPROACHES

The aim of this project is to develop an *in vitro* method for detection of sensitizing compounds (SC) and thereby reduce the number of living animals in predictive sensitization procedures. 3/4 teams focused on *in vitro* hapten processing, antigen presentation and T-cell stimulation in the murine, human and guinea-pig cellular system respectively. The emphasis on morphological and functional changes induced by strong to weak sensitizers versus non-sensitizers and hapten-peptide derivatives synthetized by the laboratory of Dermatochimie (Strasbourg) was studied.

RESULTS AND DISCUSSION

I. University of Strasbourg

The major problem with xenobiotics is their low solubility in culture media. Several approaches to this problem therefore have been developed.

The first was to introduce water-soluble functionnal groups far from the reactive site of the molecule, the second was to synthesize pro-drugs by introducing water-soluble groups that could subsequently be cleaved (cell metabolism, pH conditions...) to release the original hapten *in situ*. These two approaches have been applied to Pentadecyl cathecol (PDC), leading to water-soluble compounds.

Hapten-protein interactions have been shown to be one of the key steps in Allergic Contact Dermatitis. The formation of the hapten-protein complex occurs mainly through a covalent bond between the hapten and nucleophilic functionalities on proteins. It is most likely that modifications of the physico-chemical properties of allergens (moving from lipophilicity) should have consequences on the number and the nature of modified amino acids.

Nuclear magnetic resonance spectroscopy (NMR) could provide direct and rapid qualitative and semi-quantitative information about the reactive and accessible amino-acids. We have thus attempted to determine whether this could be used to analyze hapten-protein interactions. Methyl alkanesulfonates, which were recently

reported to cause severe allergic contact dermatitis in guinea pigs and mice, were used as a model substance. These compounds act as lipophilic methylating agents, transferring a methyl group to nucleophilic sites of aminoacids in proteins. We used ^{13}C-labelled sulfonates in association with ^{2}D-reversed (^{1}H, ^{13}C) correlation NMR to identify the modified amino acids on large proteins such as human serum albumin or ovalbulmin. Methyl alkanesulfonates were found to react mainly with histidine residues while an α-methylene-γ-butyrolactone, another well known hapten, was found to react mainly with lysines. This approach seems very promising for the elucidation of parameters involved in the generation of T epitopes by small reactive molecules and we are now preparing ^{13}C-labelled analogues of 5-and 6-methyl PDC (see figure below).

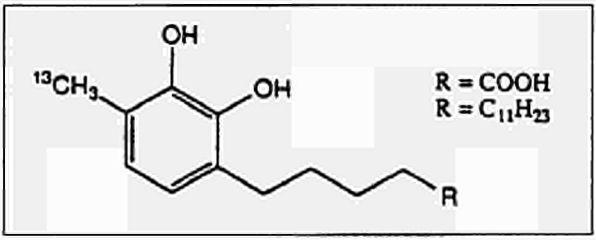


Fig. 1:
^{13}C-labelled analogues of PDC

II. Sanofi Recherche

1. Experiments with Guinea Pig Epidermal Cells (EC)

The first year, optimal conditions for the isolation and preparation of guinea pig EC were defined and optimal hapten concentrations were determined for each xenobiotic. We have analyzed the production of cytokines by the EC in suspension, after pulse stimulation with haptens. Under these conditions, results showed a high background production of cytokines certainly linked to the cell preparation.

The second year program was carried out with the determination of optimal conditions as regards the haptenization of EC in culture. The keratinocytes were cultured with feeder layer and then haptenized for 30 min with different compounds. The cytokine production was measured at various times. As specific markers or antibodies for guinea-pig are not available, measurements were limited to two cytokines, for which bioassay are possible: IL-6 (B9 line cell assay); TNF α-(L-929 cytolytic assay). PGE2 secretion was studied using a radioimmunoassay. Results showed that cytokine and PGE2 production was not specifically enhanced by sensitizers versus irritants.

2. Experiments with Human EC

In view of the problems encountered with the guinea-pig cell system (lack of specific markers or antibodies for guinea-pig), we focused, for the last year, on the studies in human EC. During the initial phase of Contact Dermatitis, the Langerhans cells (LC) play a pivotal role in antigen processing and presentation to the immunocompetent T cells. It is now recognized that this process involves Receptor Mediated Endocytosis (RME). We studied the endocytosis of class II molecules in freshly isolated human LC. Two techniques were used to analyse HLA-DR molecule endocytosis by freshly isolated LC: flow cytometry and confocal microscopy.

LC were labelled by indirect immunofluorescence with monoclonal antibodies (MoAb) directed against HLA-DR molecules and then with FITC-conjugated goat anti-mouse IgG polyclonal Ab for flow cytometric analysis or with Cy3-conjugated goat sheep anti-mouse IgG for confocal microscopic analysis.

In preliminary experiments, we determined the endocytosis kinetics of HLA-DR molecules by human LC. The cell suspensions were incubated at 37°C for 15, 30, 45, 60 and 90 min. The dead cells were labelled with propidium iodide (PI) and the results were analysed by flow cytometry. This analysis was performed on HLA-DR positive (HLA-DR +), PI negative cells (PI-). Under these conditions, results showed a time-dependent decrease in fluorescence intensity, though no increase in cell mortality was observed (figure 2). This decrease was certainly due to the HLA-DR endocytosis in acid compartments, such as lysosomes in which FITC is degraded. In subsequent experiments, we studied the effect of different compounds such as DNFB, $NiSO_4$, DNBS, or SDS on HLA-DR molecule endocytosis by LC. The human EC were incubated with these compounds at 37°C for 30 min. The flow cytometric analysis showed a dose-dependent decrease in fluorescence intensity of HLA-DR + cells (figure 3). There was no significant difference between sensitizers and irritants.

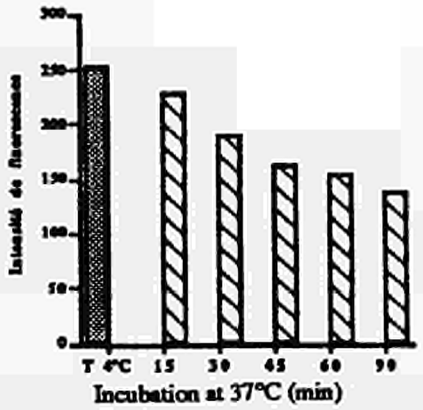


Fig. 2: HLA-DR expression by human LC after incubation at 37°C

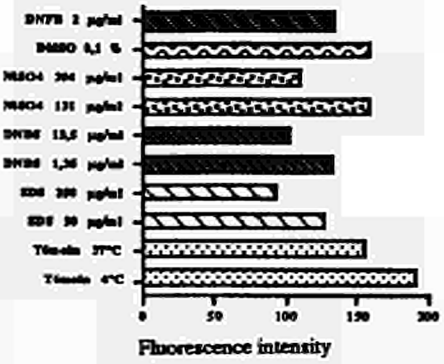


Fig. 3: Effect of sensitizers or irritants on HLA-DR expression by human LC

Confocal microscopic analysis was performed after fixation of EC with 1% Paraformaldehyde (PFA). This analysis showed different stain pattern depending on the incubation period:
T = 0 min continuous peripheral staining,
T = 15 min discontinuous peripheral staining,
T = 30 min discontinuous peripheral staining or vesicular intracellular staining,
T = 60 min diffuse vesicular intracellular staining.
After stimulation with sensitizers, a more intense fluorescence and dense intracellular staining were detected compared to the control sample staining. The samples incubated with the irritant (SDS) presented a pattern similar to that of the control samples.

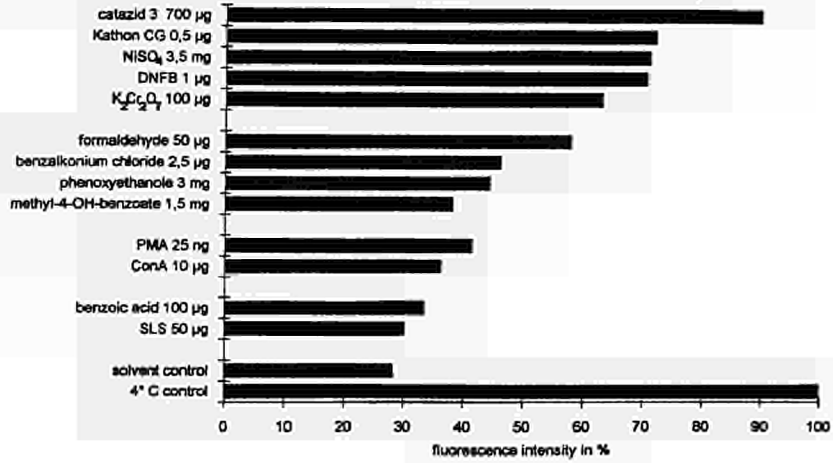
Flow cytometry supplies quantitative information on HLA-DR endocytosis by human LC and could be used as a technique to screen compounds. Confocal microscopy supplies qualitative information and characterizes the type of staining. Both methods were complementary as regards our study model and can be used to study the sensitizing potential of chemical compounds.

III. University of Mainz

1. Endocytotic activation of murine epidermal LC by contact sensitizers as parameter for in vitro testing of small reactive chemicals.

Based on earlier work using as parameter the qualitative distribution of internalized gold-labelled MHC class II molecule-specific antibodies under the influence of various compounds (see former report) a flow cytometric *in vitro* test was developed. EC suspensions were labelled with a monoclonal Ab directed against MHC class II molecules and fluorochrome- or biotin-coupled second step reagents. After a short time culture in the presence of sensitizing compounds, irritants or stimulating substances, the cell suspensions were labelled with a second class II molecule-specific MoAb and corresponding second step reagents as well as PI. For untreated LC a chloroquine-sensitive, strong decrease in the mean fluorescence intensity was noticed using FITC-conjugated secondary antibodies. These data suggest that class II molecules under these conditions are internalized into an acidic compartment leading to a decrease in the FITC-emission at 530 nm. This mechanism was not significantly affected by irritants (SLS, benzoic acid) or stimulating agents (concanavalin A (Con A), phorbol 12-myristate 13-acetate (PMA)) when used in subtoxic concentrations. Stimulation with sensitizing compounds such as DNFB, Catazid 3, Kathon CG, $K_2Cr_2O_7$ or $NiSO_4$ resulted in partial maintenance of the fluorescence intensity (figure 4).

Fig. 4: Fluorescence intensity of internalized MHC-class II-specific antibodies after their uptake into stimulated or untreated LC. The intensity of LC kept at 4°C was defined as 100%. Results of a typical experiment using the highest subtoxic concentrations of the indicated compounds are shown.



This effect correlated with the increase reported previously in the stimulation index which can be calculated from the internalization pattern of labelled MHC class II molecules on the single cell level. From these data we conclude that when stimulated with contact sensitizers LC internalize cross-linked MHC class II molecules into less acidic, submembraneous compartments. The degree of modulation correlated with the allergenicity of the compound. Strong sensitizers such as DNFB, Kathon CG or Catazid 3 were highly reactive whereas irritants such as SLS and benzoic acid or stimulating but not sensitizing substances such as PMA and Con A failed to mediate a significant modulation of RME. Compounds with weak allergenicity in humans or BALB/c mice were moderately ($K_2Cr_2O_7$, formaldehyde) reactive or only reactive at very high doses ($NiSO_4$). The reactivity of the other substances tested (phenoxyethanole, methyl-4-OH-benzoate, benzalkonium chloride) was not convincing enough to characterize them as potent sensitizers in this assay. Indeed these compounds are very weak (methyl-4-OH-benzoate) or rare sensitizers (phenoxyethanole) or more often described as irritant (benzalkonium chloride).

This cytometric approach represents an objective method for fast and reliable detection of endocytotic activation of LC by contact sensitizers and should be further evaluated for its usefulness for predictive testing of new compounds.

2. Modulation of receptor-mediated endocytosis of membrane molecules by human blood-derived dendritic cells under stimulation with CS.

In an attempt to perform similar studies using human cells, we investigated the influence of contact sensitizers on endocytotic mechanisms of human blood-derived dendritic cells (DC). DC were distinguished from other MHC class II-bearing cell populations by staining for cell lineage-specific markers (CD3, CD14, CD19) which are not expressed by DC. Using similar flow cytometric principles as established for murine LC we were able to demonstrate the internalization of labeled MHC class II molecules into less acidic compartments under stimulation with contact sensitizers in comparison to untreated cells. Our present results demonstrate that cytometric evaluation of DC activation by contact sensitizers is possible and that this method should be further evaluated for predictive testing of contact sensitizers using human material.

3. Regulation of cytokine-specific mRNA in cultured keratinocytes after stimulation with CS.

The usefulness of hapten-mediated production of keratinocyte-derived cytokines as a predictive parameter for detection of haptens *in vitro* was studied by monitoring expression of IL-8 and TNFα-specific mRNA in long term cultured human keratinocytes and HaCaT cell line using Northern-blot analysis. Althought significant expression of cytokine-specific transcripts was detectable, no hapten-specific effect could be demonstrated.

IV. University of Amsterdam

1. Effect of contact allergens on the expression of membrane molecules on separate Antigen Presenting Cell (APC) in vitro

Human peripheral blood monocytes were pulsed with contact allergens or irritants, and analyzed for changes in expression of HLA-DR, ICAM, VCAM, LFA-1 and LFA-3. Upregulation of ICAM and VCAM expression was found after 16 hours incubation with $NiSO_4$ and $CuSO_4$ compared to non-treated, LPS and IFN-γ treated monocytes. These changes could not be demonstrated with DNFB and dinitrochlorobenzene (DNCB) as major contact allergens. HLA-DR, LFA-1 and LFA-3 expression was not altered by the contact allergens. In this first series of experiments, it can be concluded that the effect

of contact allergens on the expression of membrane molecules of human monocytes does not seem to be a parameter for the detection of contact allergenic compounds.

2. Effect of contact allergens on APC in human skin organ cultures.

A second parallel series of experiments was performed in collaboration with Dr. P. K. Das of the department of pathology. The effect of contact allergens: DNFB, $NISO_4$, DNCB, Oxazolone (OX), diphencyprone (DPC), potassium dichromate ($K_2Cr_2O_7$), and pentadecyl-cathecol (PDC); irritants: benzalkonium chloride (BC), nonanoic acid (NAA), propylene glycol (PG), SDS and croton oil (CO); and tolerogens: 2,3 DCNB, 2,4 DCNB on APC in human skin organ cultures were studied. Contact allergens were tested on three different skin organ cultures, while the irritants and tolerogens were tested on two different skin cultures.

MHC class II and CD1a expression in epidermis were decreased 24 and 48 hours after application of various concentrations of all contact allergens, except the contact allergen PDC (this allergen was tested three times in various concentrations, but no effects were found in epidermis) (figure 5). The irritants and tolerogens have little or no effect on the expression of CD1a and HLA-DR in epidermis, in relation to the control skin culture.

From these experiments it can be concluded that contact allergens preferentially affect LC, while the non-allergenic irritants and tolerogens have little or no effect on LC. The present findings suggest that the skin organ culture system provides a promising model for the detection of allergenicity of chemical compounds.

Fig. 5: Effect of contact allergens, irritants and tolerogens on CD1a and HLA-DR expression in epidermis

Compound	*CD1a*	*HLA-DR*
contact allergens:		
DNFB	+ + + +	+ + + +
$NiSO_4$	+ + + +	+ + + +
DNCB	+ + +	+ +
OX	+ + + +	+ + +
DPC	+ + +	+ + + +
$K_2Cr_2O_7$	+ + +	+ + +
PDC	+	–
irritants		
SDS	–	+
CO	–	–
BC	+	–
NAA	–	–
PG	+	+
tolerogens:		
2,3 DCNB	–	
2,4 DCNB	+	–
vehicles	–	–

Legends: – none; + weak; + + moderate; + + + strong; + + + + total disappearance of expression of the stated membrane markers in epidermis. Mean values from three experiments with the contact allergens and two experiments with the irritants and tolerogens.

MAJOR SCIENTIFIC BREAKTHROUGHS AND/OR INDUSTRIAL APPLICATIONS

Preliminary studies on cytokine secretion in culture supernatant of guinea-pig EC as well as on mRNA expression by cultured human keratinocytes revealed great variability and little specificity.

From the discrepancy in the effects of contact allergens on the expression of ICAM and VCAM by human monocytes, it seems that these markers are not suitable for the detection of potential contact allergenic compounds.

The model of endocytosis of class II molecules by human and murine LC or human dendritic cells revealed that tested sensitizers and irritants can modulate this endocytosis. The light and confocal microscopy showed different endocytosis patterns according to the type of compound tested.

In addition, the sensitizers studied in the model skin culture induced a decrease in the number of HLA-DR + LC in the epidermis, in contrast to irritants that had little effect on LC.

We conclude that the model of class II molecule endocytosis by LC and that of skin culture were promising models for the detection of the sensitizing potential of chemical substances.
It would be necessary to confirm these results in our laboratories and in others not included in this project to establish both methods as standardized assays. Moreover, both methods have to be prevalidated with a larger range of SC, irritants and non SC.

MAJOR COOPERATIVE LINKS

— 2 work meetings per year had been organized since the official start of the project.
— 2 meetings to prepare the ELWW booklet tiled '*in vitro* integrated approach to skin pharmaco-toxicology' with two others groups (BRIDGE PL 890177, PL 890142).
— 1 st ELWW meeting: Importance of cellular interactions in skin pharmacotoxicology (Montpellier, 1993)
— 2 meetings were organized with the teams of the others BRIDGE projects (Dublin 1992, Insbruck 1994).
— 3 participants of the project have spent some time in the other team to learn some techniques developped in their laboratories.

PUBLICATIONS

Joint publication

A. Barbier, E. Rizova, J-L. Stampf, F. Lacheretz, F. Pistoor, J. Boss, M,. Kapsenberg, D. Becker, M. Mohamadzadeh, J. Knop, S. Mabic , J-P. Lepoittevin. Development of a predictive *in vitro* test for detection of sensitizing compounds. In Rougier (Ed.). Alternative methods in toxicology, vol. 10, Mary-Ann Liebert Inc. Publishers, p. 341-351, 1994. *In vitro* skin toxicology eds Maibach, 1994, (in press).

Individual publications

D. Becker, G. Kolde, K. Reske, J. Knop. An *in vitro* test for endocytotic activation of murine epidermal Langerhans cells under the influence of contact allergens. J. Immunol. Meth., 1994, **169**, 195.
D. Becker, U. NeiB, S. NeiB, K. Reske, J. Knop. Contact allergens modulate the expression of MHC class II molecules on murine epidermal Langerhans cells by endocytotic mechanisms. J. Invest. Dermatol. , 1992, **98**, 700

D. Becker, M. Mohamadzaded, K. Reske, J. Knop. Increased level of intracellular MHC class II molecules in murine Langerhans cells following in vivo and *in vitro* administration of contact allergens. J. Invest. Dermatol. , 1992, **99**, 545

S. Mabic, C. Benezra, J.P. Lepoittevin. Direct synthesis of mono-Glycosylated catechols from glycosylacetates or imidates using BF3. OEt2 as catalyst Tetrahedron Letters, 1993, **34**, 4531-4534

S. Mabic, J.P. Lepoittevin. Efficient and safe deprotection of hindered silylated ethers using *in situ* generated HF. Tetrahedron Letters, 1994, in press.

A. Barbier, J.L. Stampf, M. Pelegrin, C. Selvi Bignon and F. Lacheretz. *In vitro* effect of sensitizing compounds. The guinea-pig cell culture model: preliminary results. 4^{e} colloque SPTC, Toulouse 14 May 1993.

D. Becker, U. Lempertz, J. Knop. Endocytic activation of dendritic cells by contact sensitizers. 1st ELWW meeting: importance of cellular interaction in the skin pharmacology. SANOFI Recherche, Montpellier 25-26 November 1993.

S. Mabic, J.P. Lepoittevin. Indentification des sites nucléophiles reactifs des protéines viv-à-vis d'un alcanesulfonate de méthyl marqué au carbone 13. 5^{e} réunion du groupe thématique magnétisme nucléaire et biologie, Toulouse, 5-8 oct. 1993.

S. Mabic, C. Benezra, J.P. Lepoittevin. Synthesis and analysis of water soluble analogs of 3-n pentadecylcatechols. 8 th European Symposium on Organic Chemistry, Barcelone, 28 August — 3 September 1993.

S. Mabic, J.P. Lepoittevin. Hapten-protein interaction studies: the use of 13C labeled molecules. 1st ELWW meeting on importance of cellular interaction in the skin pharmacology. SANOFI Recherche, Montpellier 25-26 November 1993.

F. Pistoor, A. Rambukkana, P. Das, J. Bos, M. Kapsenberg. Effect of contact allergens, irritants and tolerogens on human antigen presenting cells. 1st ELWW meeting: importance of cellular interaction in the skin pharmacology. SANOFI Recherche, Montpellier 25-26 November 1993.

F. Pistoor, A. Rambukkana, M. Kroezen, J.D. Bos, M.L Kapsenberg and P.K Das. Effect of contact allergens on human Langerhans cells. 12 th meeting of the European Research Group on Experimental Contact dermatitis. Stockholm, Sweden: 6-7 May 1994.

E. Rizova, A. Barbier, P. Carayon, L. Michel, M. Pelegrin, C. Broto and F. Lacheretz. Contact allergens and Langerhans cells: early events of antigen processing. 1st ELWW meeting: importance of cellular interaction in the skin pharmacology. SANOFI Recherche, Montpellier 25-26 November 1993.

E. Rizova, P. Carayon, L. Michel, A. Barbier, M. Pelegrin, C. Broto, F. Lacheretz and L. Dubertret. Influence des allergenes de contact sur l'endocytose des molécules HLA-DR par les cellules de Langerhans humaines. 4^{e} Congrès SPTC, Paris 16-18 March 1994.

E. Rizova, A. Barbier, P. Carayon, L. Michel, M. Pelegrin, C. Broto and F. Lacheretz. Effect of contact allergens on HLA-DR molecule endocytosis by human Langerhans cells. 12 th Meeting of the European research Group on Experimental Contact Dermatitis. Stockholm, Sweden, 6-7 May 1994.

Pharmacology and toxicology of differentiated cell types, their cell-cell and cell-matrix interactions in an *in vitro* reconstructed human skin model (BIOT CT-900193)

PARTICIPANTS:

1. B. COULOMB/L. DUBERTRET, INSERM U 312, Créteil, F
2. C. LAPIERE, Univ. Liège, Tour de Pathologie, Sart-Tilman, B
3. Th. KRIEG, Lab. Dermatology, Univ. Köln, Köln, D
4. N.E. FUSENIG, DKFZ, Heidelberg, D
5. P. GIACOMONI, L'Oréal, Chevilly-Larne, F

BACKGROUND INFORMATION

Human skin was the first organ to be reconstructed *in vitro* and is likely to provide a predictive system for evaluating drug efficacy and toxicity, avoiding animal experimentation. Organogenesis *in vitro* for pharmacological studies is a very new and promising field of investigation.

Culturing cells inclose contact with their physiological matrix molecules and with cell types usually adjacent *in vivo*, cells have been shown to communicate and differentiate in this skin equivalent. These cell-matrix-cell interactions greatly modify the response to pharmaco-toxicological agents, resembling the situation *in vivo* and demonstrating that some pharmaceutical agents operate on the cell-cell and cell-matrix communication system.

OBJECTIVES AND PRIMARY APPROACHES

The objective of this program was to reproduce physiological or pathophysiological cellular regulations using *in vitro* reconstructed human living skins to develop predictive pharmaco-toxicological models.

By mimicing *in vitro* cell-matrix and cell-cell interactions of skin, we could reproduce the expression of normal skin functions and mimic defined pathological situations. The living human skin equivalent models were completed with different cell types (normal or pathological) or by adding other components of the extracellular matrix, and new parameters for analysis have been defined. We also applied model substances used in dermatology in order to check the predictive value of these models.

PARTICIPANT No.1:

During this program we could demonstrate that cell-matrix or dermal-epidermal interactions could be determinant for the cellular response to growth factors or pharmacological agents, and permit to define fibroblast phenotype. In addition, we could define culture conditions permitting to be more predictive to the *in vivo* situation in psoriatic patients.

Fibroblast-matrix interactions: IL-1 and PDGF are known to stimulate fibroblasts growth in monolayer. But in the presence of a collagen matrix, IL-1 or PDGF become less efficient on cell growth and, while IL-1 inhibits fibroblasts contraction, PDGF stimulates it. This suggests an important regulatory role of these two factors during the wound healing process. Not only the presence of a collagen matrix

influence cell response but also the 'maturation' of the fibroblasts within this matrix, as shown by a progressive loss of sensitivity of fibroblasts to the growth stimulation by aFGF.

Pharmacology and predictiveness: We previously shown the influence of fibroblasts on keratinocytes growth and more particularily their ability to modulate the effects of pharmacological agents such as retinoids. But if fibroblasts are involved in dermal-epidermal interactions, serum factors can also be of importance on the cellular response to drugs and have to be taken into account to make skin equivalent models more predictive. As an example, the presence of serum from psoriatic patients is necessary to get the response observed on the epidermis of these patients after treatment by acitretin.

The re-establishment of cell-matrix or cell-cell interactions permits also to discriminate fibroblasts of different origins according to the different phenotypes they can reveal.

Using a new PKC inhibitor, GF109203X, we could show that PKC is involved differentially in human fibroblasts and keratinocytes growth and in their response to growth factors. The specicity of this inhibitor for PKC in fibroblasts and keratinocytes, make the GF109203X a relevant tool for determination of PKC involvement.

PARTICIPANT No. 2:

The reconstruction *in vitro* of organs, such as a dermal-equivalent, using human cells, normal or pathological, amplified in culture and associated with defined extracellular matrix macromolecules has allowed to demonstrate that the chemical and mechanical signals delivered by the matrix to the cells resulted in the regulation of their phenotype. It is expressed by modifications of the proliferative and biosynthetic programme, changes in responsiveness to cytokines, inductions of differentiated morphotype or invasive behaviour. This concept has been validated for fibroblasts of various origins (skin), vascular endothelial and smooth muscle cells, granulation tissue myofibroblasts cultured within a support of instertitial collagen as well as for cancer cells and a reconstructed basement membrane. These experimental models further allowed to associate two types of cells (keratinocytes-fibroblasts, tumoral cell-fibroblasts) and to analyze the cell to cell exchange of information. In parallel, these reconstituted organs were used as alternative models to animal experimentation in pharmaco-toxicological studies. The mechanisms operating in the transduction of the information from the matrix to the cell, from the cell to the matrix and from cell to cell and the resulting intracellular signalling become understood. Pharmacological compounds, growth factors, physiological mediators and physical agents as electric and magnetic fields, were used for validating this exchange of information as potential targets for *in vitro* pharmaco-toxicology.

PARTICIPANT No. 3:

The ultimate aim was to combine dermal cells of different lineages with constituents of the extracellular matrix to form a dermis-equivalent, which can serve to be epidermalized by keratinocytes of normal or diseased skin.

Different dermal cells, e.g. fibroblasts derived from skin, and endothelial cells derived from capillaries or umbilical cords, displayed a different, cell-type specific

response to the matrix environment, however, they are equally well suited to promote the proliferation of keratinocytes. Different proliferative capacities were obtained using fibroblasts derived from healthy as compared to fibrotic skin.

Cellular phenotype and biosynthetic capacity was found to be strongly influenced by different matrix constituents. Previous work had focused on the use of the major matrix components, collagens I and III. Addition of glycosaminoglycans elicited cell-type specific responses, especially in endothelial cells.

The communication between cells, and cells and matrix, is mediated by integrins. Expression of different matrix receptors was studied in the three-dimensional model, and studies attempting to block receptor function were conducted in order to elucidate their biological function. The biochemical nature of signals which are transduced by integrins from the environment to the cytoplasm and into the nucleus was analyzed and found to involve protein phosphorylation on tyrosine and the action of phospholipase C.

PARTICIPANT No. 4:

Our studies were focused on

(i) the standardization of the organotypic keratinocytes coculture model,

(ii) the molecular analysis of factors interacting between keratinocytes and fibroblasts or endothelial cells,

(iii) the interaction of immortalized keratinocytes (HaCaT) with mesenchymal cells *in vitro*.

The major component of the organotypic culture model such as keratinocytes, fibroblasts, endothelial cells and collagen matrix could be standardized in a way so that an epidermis-like structured tissue could be routinely achieved. By modifying matrix holding devices, the stability of such cultures could be improved. Moreover, we recently succeeded in developping a serum-free and factor defined culture medium in which such organotypic coculture could be produced exhibiting even improved structure and differentiation phenotypes as compared to serum-containing medium. This model with a compact stratum corneum, normalized proliferation and differentiation and prolonged survival *in vitro* is now used for pharmacological studies.

The interplay between keratinocytes and mesenchymal cells responsible for tissue reconstruction in such organotypic cultures occurs via complex interactions by diffusible paracrine acting factors. We could detect a novel mechanism by which such factors are reciprocally induced in the cocultured cells. Thus, keratinocytes and fibroblasts mutually stimulate their proliferation via Interleukine-1α expressed by keratinocytes leading to upregulation of IL-1α receptor in fibroblasts followed by the increased expression of KGK, IL-6, and GM-CSF as well as collagenase in fibroblasts. On the other hand, TGFα, an autocrine acting keratinocyte growth factor, was not regulated by mesenchymal cells.

The use of a permanent cell line expressing most of the normal keratinocyte functions could be a considerable advantage in using organotypic cultures for routine assays. The spontaneously immortalized HaCaT cell line which exhibited rather normal differentiation capacity in transplants *in vivo* was a good candidate for this kind of studies. Furthermore, we selected a highly differentiated subclone which

formed stratifying epithelia in organotypic culture assays with rather normal localization of differentiation parameters. Their growth behaviour was similarly dependent on mesenchymal cells but orderly structure epithelia were not formed in serum nor serum-free culture conditions, but can be further normalized by addition of different growth factors. Interactions of HaCaT cells with the collagen matrix via integrins indicated a great similarity with normal keratinocytes.

PARTICIPANT No. 5:

This group have analysed the requirements for growing normal human melanocytes in culture, alone or in the presence of normal human keratinocytes, and set up methodologies for the quantification of intra-cellular melanins.

One of the aims of this work was to study if a correlation exist between the pigmentation which can be obtained *in vivo* and an accumulation of melanin within cultured melanocytes. As a matter of fact, repeated UV-B irradiations induce a phototype-dependent pigmentation in humans within few days, and UV-A combined with psoralens are strong inducers of pigmentation.
In vitro, we have observed that repeated low doses of UV-B irradiations (2,4 mJ/cm^2 per day) at 285 nm usually double the intra-cellular melanin content in cultured melanocytes within four days after the first irradiation. The amount of intra-cellular melanin depends on donors and culture conditions (presence of serum,...), but is generally between 9 and 58 pg/cell.

We have observed that repeated daily UV-A irradiations (7 J/cm^2 at 365 nm for four days) do not increase the intra-cellular level of melanin. When irradiations are performed in the presence of 8 methoxy psoralen (5-20 μg/ml) the intra-cellular melanin is multiplied by three within two days after the first irradiation and melanocytes undergo morphological modifications (melanin is secreted and forms aggregates). Angelicin or Khellin or 5-methoxy psoralen also increase the content of melanin (about 50%).

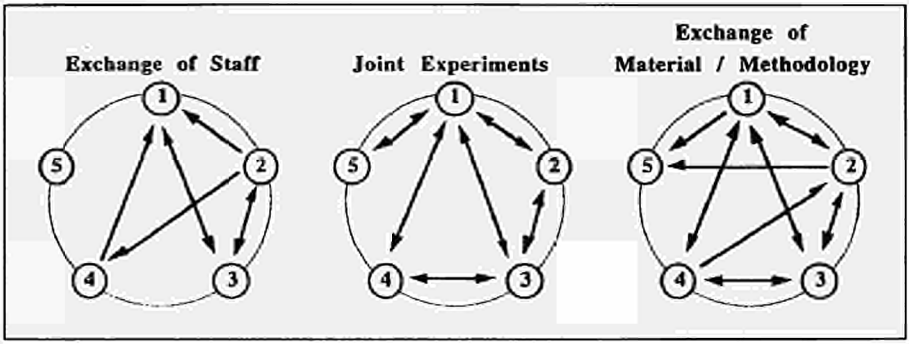
We also attempt to combined melanocyte-keratinocyte coculture on dermal equivalent (containing fibroblasts or endothelial cells). In these conditions, keratinocytes and melanocytes fail to mix and thus do not display an organized structure. Melanocytes tend to cluster, only occasionnaly melanocytic dendrites heading towards, or interacting with, keratinocytes are observed.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The present state of these *in vitro* human skin equivalent models represents a close approximation to human skin enabling further studies on regulatory mechanisms important for tissue homeostasis and diseased states. The established models allow analysis of the influence of external factors on a multicellular organ system promissing relevant data for the *in vivo* situation. More work is required to further standardization, but preliminary test series have already been started in collaboration with industrial companies for its potential to replace animal experiments.

MAJOR COOPERATIVE LINKS

Within this BRIDGE Project:



With other BRIDGE Projects (A. Barbier, Sanofi & E. Honikman-Leban, L'Oréal):

— Creation of an ELWW: 'In the field of *in vitro* integrated approach to skin pharmaco-toxicology'
— Organization of a meeting 'Importance of cellular interactions in the skin pharmacotoxicology' in Montpellier (25-26 Nov. 1993)

PUBLICATIONS

Joint publications

Scharffetter-Kochanek K., Klein C.E., Heinen G., Mauch C., Schaefer T., Adelmann-Grill B., Goerz G., Fusenig N., Krieg T., Plewig G. Migration of human keratinocyte cell line HACAT to interstitial collagen type I is mediated by a2β1 integrin receptor. J Invest Dermatol 98: 3-11, 1992 (participants 3 & 4).

Lapière ChM and Krieg Th editors, Connective Tissue Diseases of the Skin, Marcel Dekker, Inc. Publ., New York, 408 pages, 1993 (participants 2 & 3).

Eckes B., Krieg T., Nusgens B.V., Lapière Ch.M. Potential applications of an *in vitro* reconstituted matrix for the understanding of skin biology and pathology. J Invest Dermatol [submitted] 1994 (participants 2 & 3).

Smola H., Thiekötter G., Stark H.J., Breitkreutz D., Hafner G., Krieg Th., Fusenig N.E. Expression of vimentin in normal human keratinocytes associated with adhesional stress. [submitted] (participants 3 & 4).

Smola H., Thiekötter G., Stark H.J., Mirancea N., Krieg Th., Fusenig N.E. Expression of basement membrane components by keratinocytes and fibroblasts in organotypic cultures. [in preparation] (participants 3 & 4).

ELWW Blue book: In the field of *in vitro* integrated approach to skin pharmaco-toxicology.

Individual publications.

Participant No. 1

Dubertret L., Bruner-Ferber F., Misiti J., Thomas K.A. and Dubertret M.L. Activities of human acidic fibroblasts growth factor in an *in vivo* dermal equivalent model. J Invest Dermatol **97**: 793-798, 1991.

Bertaux B., Hornebeck W., Eisen A.Z. and Dubertret L. Growth stimulation of human keratinocytes by tissue inhibitor of metalloproteinase. J. Invest Dermatol **97**: 679-685, 1991.

Coulomb B. and Dubertret L. Production and use of skin cell cultures for the *in vitro* testing of chemicals and pharmaceuticals. Animal Cell Biotechnology, **5**: 123-149, 1992.

Coulomb B. and Dubertret L. Interactions dermo-épidermiques et pharmacologie cutanée. Pathol. Biol. **40**: 139-146, 1992.

Coulomb B., Dubertret L. Three-dimensional reconstructed human skin for *in vitro* pharmaco-toxicology. In: Human cells in *in vitro* pharmaco-toxicology. Rogiers V., Sonck W., Shepard E., Vercruysse A. eds; Vubpress, Brussels, 27-44, 1993.

Coulomb B. Culture des cellules de la peau: Nouvelles approches. Nouvelles Dermatol, **12**: 438-442, 1993.
Coulomb B. and Dubertret L. *In vitro* reconstructed skin for grafting in human. In: Connective Tissue Diseases of the skin. Lapière Ch., Krieg T. eds; Marcel Dekker, N.Y., 349-360, 1993.
Le Panse R., Coulomb B., Mitev V., Bouchard B., Lebreton C. and Dubertret L. Differential modulation of human fibroblast and keratinocyte growth by the protein kinase C inhibitor, GF 109203X. Molecular Pharmacology, **46** (volume) (in press), 1994.
Le Panse R., Mitev V., Lebreton C. and Coulomb B. Modulation of epidermal growth factor (EGF) and keratinocyte growth factor (KGF) effects on human keratinocyte growth by protein kinase C inhibitor, GF 109203X. (Biochem. Biophys. Res. Commun. (in press), 1994.
Sanquer S., Coulomb B., Lebreton C. and Dubertret L. Effect of acitretin on epidermal growth in a human skin model containing fibroblasts and serum from psoriatic patients (submitted).
Friteau L., Coulomb B., Lebreton C., Damais C. and Dubertret L. Effects of Interleukin-1b and platelet-derived growth factor-BB on human fibroblast behaviour in the absence or presence of a collagen matrix. (submitted).

Participant No. 2

Delvoye P., Wiliquet P., Lévèque J.L., Nusgens B.V. and Lapière Ch.M. Measurement of mechanical forces generated by skin fibroblasts embedded in a three-dimensional collagen gel. J Invest Dermatol **97**: 898-902, 1991.
Colige A., Roujeau J.C., De la Rocque F., Nusgens B. and Lapière Ch.M. Abnormal gene expression in skin fibroblasts from a Hutchinson-Gilford patient. Lab Invest **64**: 799-806, 1991a.
Lambert C.A., Soudant E.P., Nusgens B.V. and Lapière Ch.M. Pretranslational regulation of extracellular matrix macromolecules and collagenase expression in fibroblasts by mechanical forces. Lab Invest **66**: 444-451, 1992.
Colige A., Nusgens B. and Lapière Ch.M. Altered response of progeria fibroblasts to epidermal growth factor. J. Cell Sci. **100**: 649-655, 1991b.
Noël A.C., Calle A., Emonard H.P., Nusgens B.V., Simar L., Foidart J.B., Lapière Ch.M. and Foidart J.M. Invasion of reconstituted basement membrane matrix is not correlated to the malignant metastatic cell phenotype. Cancer Res **51**: 405-414, 1991.
Colige A.C., Lambert C.A., Nusgens B.V. and Lapière Ch.M. Effect of Cell-Cell and Cell-Matrix Interactions on the Response of Fibroblasts to Epidermal Growth Factor *In vitro* — Expression of Collagen Type-I, Collagenase, Stromelysin and Tissue Inhibitor of Metalloproteinases. Biochem. J. **285**: 215-221, 1992.
Lapière Ch.M. and Nusgens B.V. La matrice extracellulaire et sa régulation. Pathol. Biol. **40**: 133-138, 1992.
Noël A., Munaut C., Boulvain, Calberg Bacq C.M., Lambert C.A., Nusgens B.V., Lapière Ch.M. and Foidart J.M. Modulation of collagen and fibronectin synthesis in fibroblasts by normal and malignant cells. J Cell Biochem. **48**: 150-161, 1992.
Noël A., Munaut C., Nusgens B.V., Foidart J.M. and Lapière Ch.M. The Stimulation of Fibroblasts Collagen Synthesis by Neoplastic Cells Is Modulated by the Extracellular Matrix. Matrix **12**: 213-220, 1992.
Lambert C.A., Lefèbvre P.Y., Nusgens B.V. and Lapière Ch.M. Modulation of expression of endogenous collagenase and collagen genes by electroporation. Possible involvement of $Ca2^+$ and protein kinase C. Biochem. J. **290**: 135-138, 1993.
Nusgens B.V. Cell-Cell and Cell-Matrix Interactions. In: Connective Tissue Diseases of the Skin, edited by Lapière, Ch.M. and Krieg, Th. New York: Marcel Dekker, Inc., 147-156, 1993.
Noël A., Nusgens B.V., Lapière Ch.M and Foidart J.M. Interaction between tumoral MCF7 cells and fibroblasts on matrigel and purified laminin. Matrix **13**: 267-273, 1993.
Rompen E.H., Kohl, J., Nusgens B.V. and Lapière Ch.M. Kinetic aspects of gingival and periodontal ligament fibroblasts attachment to surface-conditioned dentin. J Dental Res. **72**: 607-612, 1993.
Lambert C.A., Lefèbvre P.Y., Deroanne C., Nusgens B.V. and Lapière Ch.M. Mechanical tension regulates the phenotype of cells cultured in a collagen gel. In: NATO Advanced Study Institute of Biomechanics of Active Movement and Division of Cells', Publ. Springer-Verlag, 519-524, 1994.

Antoine N., Greimers R., Deroanne C., Kusaka M., Heinen E., Simar L.J. and Castronovo V. AGM-1470, a potent angiogenesis inhibitor, prevents the entry of normal, but not transformed, endothelial cells into the G1 phase of the cell cycle. Cancer Res [in press], 1994.

Participant No. 3

Klein C.E., Dressel D., Steinmayer T., Mauch C., Eckes B., Krieg T., Bankert R.B., Weber L. Integrin α2β1 is upregulated in fibroblasts and highly aggressive melanoma cells in three-dimensional collagen lattices and mediates the reorganization of collagen I fibrils. J Cell Biol **115**: 1427-1436, 1991.
Heckmann M., Aumailley M., Chu M.L., Timpl R., Krieg T. Effect of transforming growth factor-β on collagen VI expression in human dermal fibroblasts. FEBS Lett **310**: 79-82, 1992.
Eckes B., Hunzelmann N., Ziegler-Heitbrock H.W., Urbanski A., Luger T., Krieg T., Mauch C. Interleukin-6 expression by fibroblasts grown in three-dimensional gel cultures. FEBS Lett. **298**: 229-232, 1992.
Mauch C., Kozlowska E., Eckes B., Krieg T. Altered regulation of collagen metabolism in scleroderma fibroblasts grown within three-dimensional collagen gels. Exp. Dermatol 1: 185-190, 1992.
Mauch C., Eckes B., Hunzelmann N., Oono T., Kozlowska E., Krieg T. Control of fibrosis in systemic scleroderma. J. Invest. Dermatol. **100**: 92s-96s, 1993.
Oono T., Specks U., Eckes B., Majewski S., Hunzelmann N., Timpl R., Krieg T. Expression of type VI collagen mRNA during wound healing. J Invest Dermatol **100**: 329-334, 1993.
Eckes B., Mauch C., Hüppe G., Krieg T. Downregulation of collagen synthesis in fibroblasts within three-dimensional collagen lattices involves transcriptional and posttranscriptional mechanisms. FEBS Lett. **318**: 129-133, 1993.
Eckes B., Krieg T. *In vitro* systems to study fibroblast functions. In: Human cells in *in vitro* pharmaco-toxicology. V. Rogier, W. Sonck, E. Shephard, A. Vercruysse (eds), VUB Press Brussels, 11-26, 1993.
Roeckel D., Krieg T. Three-dimensional contact with type I collagen mediates tyrosine phosphorylation in primary human fibroblasts. Exp Cell Res. **211**: 42-48, 1994.

Participant No. 4

Machenzie I., Rittman G., Bohnert A., Breitkreutz D., Fusenig N.E. Influence of connective tissues on the *in vitro* growth and differentiation of murine epidermis. Epith Cell Biol. **2**: 107-119, 1993.
Limat A., Breitkreutz D., Thiekötter G., Noser F., Hunziker T., Braathen L.R., Fusenig N.E. Phenotypic modulation of human hair matrix cells (trichocytes) by environmental influence *in vitro* and *in vivo*. Epith Cell Biol. **2**: 55-65, 1993.
Smola H., Thiekötter G., Fusenig N.E. Mutual induction of growth factor gene expression by epidermal-dermal cell interaction. J Cell Biol. **122**: 417-429, 1993.
Breitkreutz D., Stark H.J., Plein P., Baur M., Fusenig N.E. Differential modulation of epidermal keratinization in immortalized (HaCaT) and tumorigenic human skin keratinocytes (HaCaT-ras) by retinoic acid and extracellular Ca^{2+}. Differentiation **54**: 201-217, 1993.
Reichrath J, Hügel U., Klaus G., Fusenig N.E. Modulation of 1,25-Dihydroxyvitamin D3 receptor expression in HaCaT keratinocytes. In: Cell and Tissue Culture Models in Dermatological Research, Bernd A, Bereiter-Hahn J, Hevert F, Holzmann H (eds), Springer Verlag, Berlin, Heidelberg, 37-43, 1993.
Limat A., Hunziker T., Breitkreutz D., Fusenig N.E., Braathen L.R. Organotypic cocultures as models to study cell-cell and cell-matrix interactions of human hair follicle cells. Skin Pharmacol. **7**: 47-54, 1994.
Fusenig N. Epithelial-mesenchymal interactions regulate keratinocyte growth and differentiation *in vitro*. In Keratinocyte Handbook. Leigh I, Watt F and Lane B editors. Cambridge University Press, 71-74, 1993.
Limat A., Breitkreutz D., Hunziker T., Klein C.E., Noser F., Fusenig N.E., Braathen L.R. Outer roots sheath (ORS) cells organize into epidermoid cyst-like spheroids when cultured inside Matrigel: a light-mocroscopic and immunohistological comparison between human ORS cells ans interfollicular keratinocytes. Cell Tissue Res. **275**: 169-176, 1994.
Fusenig N.E., Limat A., Stark H.J., Breitkreutz D. Modulation of the differentiated phenotype of keratinocytes of the hair follicle and from epidermis. J Dermatol. Science, (in press), 1994.

Establishment of immortal, differentiated hepatocyte lines from transgenic mice harboring SV40-T-ag sequences and their use in toxicology and molecular genetics ((BIOT CT-900189)

COORDINATOR:
D. PAUL, Fraunhofer Inst., Hannover, D
PARTICIPANTS:
C. GUILLOUZO, INSERM U49, Rennes, F

BACKGROUND INFORMATION

Animal cells from numerous normal tissues can be maintained in culture. However, only few cell types can be cultured as non-tumorigenic proliferating populations for extended periods unless they have been established as immortal cell lines. The most commonly used cell lines are rodent 'fibroblastoid' cell types that can readily be grown and selected for permanent growth in culture, but they do generally not maintain the specific differentiated phenotype of their tissue of origin. In contrast, cells from normal human tissues have been established in only a few instances as established cell lines. However, in some cases human cells been have successfully been cultured as diploid cell strains that do not display permanent growth potential but show a finite life span.

In contrast, a large number of animal cells has been obtained. Few have been described that maintain normal patterns of specific functions of the tissue from which they were prepared. Even fewer cell lines were derived from epithelial cell types, which are among the most important cells in cancer research because carcinomas, the most frequent tumor type in man, are derived from them. More recently, such cell lines could be established following immortalization with genes of transforming viruses.

Although cell type specific functions have frequently been found to be expressed in these lines and have therefore been useful in biomedical research, they generally do not faithfully reflect the properties of the normal cell types in tissues from which they were derived.

In view of these difficulties, the availability of truly differentiated cell lines expressing the complete set of cell type-specific functions would be of great importance in these areas. Of particular interest are culture systems such as immortal hepatocyte lines that are useful for numerous applications (Table I).

For example, differentiated hepatocyte lines in culture are in great demand that express normal drug metabolizing enzyme systems that lend themselves for the development of toxicological short-term test systems to identify directly or indirectly acting mutagenic compounds. Differentiated, immortal but non-transformed cell lines that acquire tumorigenic properties upon long-term cultivation are of interest to study genetic events that accompany transformation. Furthermore, differentiated hepatocyte lines are needed to investigate the basic biology of viruses such as Hepatitis B-virus (HBV), which specifically infects, expresses its genes and replicates in hepatocytes. Finally, such cell lines are in great need for the production of recombinant human proteins with pharmacological activities that are not accessible by other means. For example, cell lines are of interest that, after appropriate genetic manipulations, maintain stable expression of functions that are required for introducing normal post-translational modifications into the recombi-

nant glycoproteins to be produced. Even though human cell lines would presumably be the optimal choice for biomedical research and for biotechnological processes, technical difficulties faced at present do not appear to make it likely that the required, permanently growing immortal human cell lines will be generated soon. Therefore, rodent cells, which are well-known to readily acquire the immortal phenotype either spontaneously or induced by specific genes will continue to play a central role in studies in the fields of cancer research, toxicology and biotechnology. However, it remains a difficult problem to generate immortal, truly differentiated cell lines that faithfully express normal patterns of functional properties that are characteristic of normal tissues from which they were derived.

TABLE I

Use of immortal, differentiated hepatocyte lines in biomedical research

Research field	*Specific applications*
1) Toxicology	Short-term mutagenicity tests
2) Cancer research	Genetic events leading to tumorigenic HCC cells
	Growth control of hepatocytes
3) Virology	Biology of HBV
4) Biotechnolgy	Production of recombinant of human proteins with pharmacological activity
5) Growth control	Regulation of the cell cycle
	Growth factor requirements

OBJECTIVES AND PRIMARY APPROACHES

The goal of this project was to obtain diploid, immortal hepatocytelines derived from adult transgenic mouse strains bearing metallithionein (MTq-I)-driven SV40T-antigen constructs for the development of *in vitro* systems for the identification of potentially mutagenic agents. The cell lines were also employed for transfecting Hepatitis B virus (HBV) DNA to study HBV gene expresion and replication *in vitro*.

RESULTS AND DISCUSSION

a) Immortalization of hepatocytes

Rodent hepatocytes in primary culture maintain their differentiated functions for only a few days and cannot be subcultured without loss of the hepatic phenotype. To establish long-term hepatocyte cultures, three main approaches were taken:

(a) modification of the culture medium to improve survival of hepatocytes in primary culture and to stabilize differentiated liver functions in serum-free cultures,
(b) establishment of cocultures of hepatocytes with other liver cell types, and
(c) immortalization of hepatocytes in primary culture by infection with SV40 virus to establish multiplying, differentiated lines.

In attempts to maintain hepatocytes in primary culture for extended periods, unconventional complex culture media were developed that permit to maintain hepatocytes in the presence of hepatotrophic growth factors without serum. Medium MX-83 was shown to stabilize the expression of liver-specific functions for at least 14 days as determined by analyses using either Northern hybridization or nuclear run-on transcription assays. In contrast, hepatocytes cultured in conven-

tional media with serum tend to dedifferentiate within 1-3 days, while hepatocytes in coculture are stably differentiated but fail to divide. However, in spite of the improvement that permits the maintenance of the hepatocyte phenotype in primary culture in MX-83 medium, it has not been possible to passage cells into secondary culture without decay of a numerous hepatic functions.

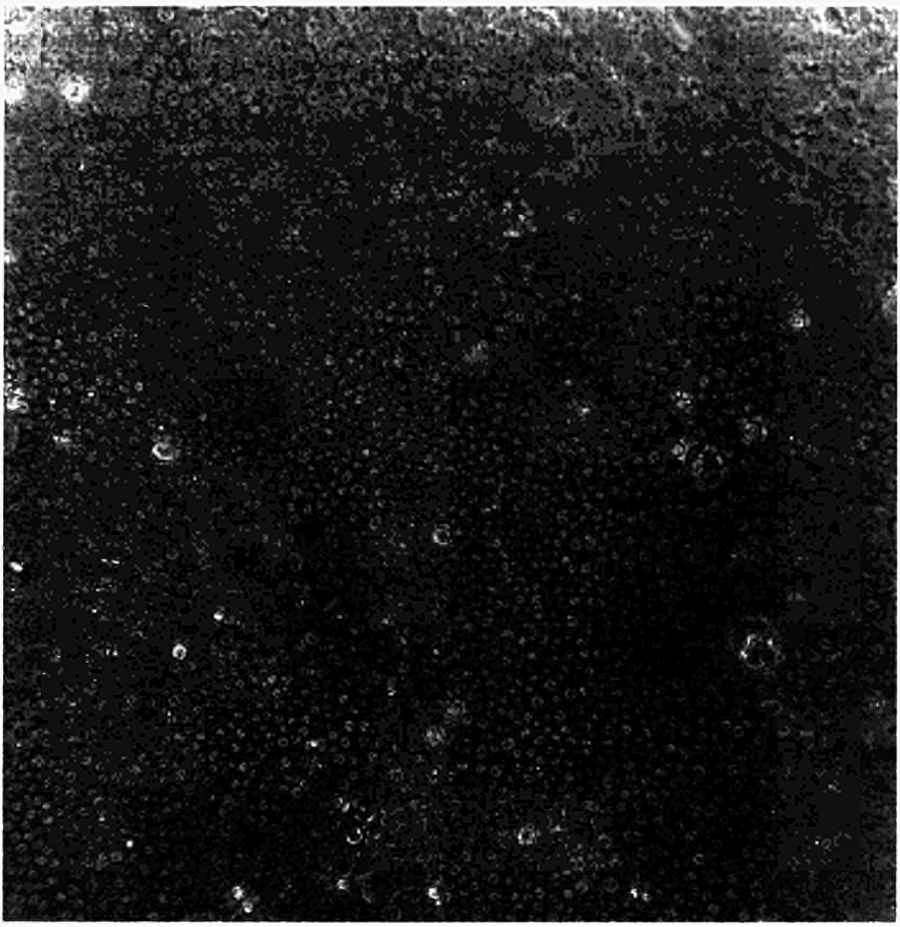


Fig. 1. Photomicrograph of fetal hepatocyte line SV-FMH-12 immortalized by SV40 virus. Phase contrast, X 200.

To establish multiplying hepatocytes in culture, primary fetal, newborn or adult liver cells were placed in culture in chemically defined MX-83 medium free of arginine in the presence of epidermal growth factor (EGF), insulin and hydrocortisone and infected with SV40 virus. After 2-3 weeks, when the viral genome is stably integrated, colonies started to grow out that maintained the normal

hepatocyte-like morphology (Fig. 1). Cells contain the viral genome stably integrated in their DNA. They multiply in response to the added growth factors and require tight cell-cell contacts with neighboring cells for growth and can therefore not be cloned. During early passages, cells maintain liver-specific functions as determined by Northern hybridization analyses. However, after culture periods of about 15-20 passages these hepatocyte lines tend to dedifferentiate, however their growth appeared to be unrestricted. Presumably, these lines tend to switch off the expression of liver specific genes rather early after they were established, because during the process of the integration of the viral genome, i.e. prior to stable expression of T antigen and to immortalization, the maintenance of numerous hepatic functions becomes unstable and cannot be subsequently re-stabilized.

In attempts to establish growing hepatocyte lines that truly maintain their differentiated functions for extended periods, a new approach was developed, which involves placing hepatocytes in culture that already contain the immortalizing gene(s) stably integrated, i.e. hepatocytes from transgenic mice bearing an appropriate transgene. Initially, transgenic mice were employed that contain the SV-GH-202 construct, which inevitably develop hepatocellular carcinoma (HCC) at the age of about 6 months. When hepatocytes prepared from livers of these mice prior to tumor development were placed into primary culture, colonies grew out after a few days and hepatocyte lines were readily established. Lines were prepared from fetal, newborn and young adult livers of transgenic SV-202 mice, which multiply in response to EGF and insulin (see below). Resulting lines are immortal and non-tumorigenic upon s.c. injection into Our observations indicate that the transforming genes of SV40 virus did not directly render normal hepatocytes tumorigenic. They suggested that T-ag expressing hepatocytes presumably accumulated specific genetic alterations that cause alterations in the transcriptional control of cellular genes involved in the control of the cell cycle, which eventually resulted in the appearance of HCC cells. Attempts were made to define a preliminary sequence of events that accompany the development in culture of untransformed transgenic hepatocytes towards HCC cells. Immediately after cultivation, the cells require EGF plus insulin for growth. After about 8-12 passages, cells abrogate their requirement for added EGF and cells multiplied in response to insulin (or to IGF I). During the following 35-40 passages cells abrogate the requirement for added insulin, which results in the propagation of autonomous cells that no longer require growth factors for multiplication. Cells usually remain strictly diploid until about passage 50. Alterations in the number of chromosomes start to become apparent at the time when cultures require low levels of growth factor (either 300 — 500 ng/ml of insulin or 300 pg/ml IGF I).

Cells grow in tight monolayer colonies (similar to those shown in Fig. 1) and cell-cell contacts are essential requirements for growth. Cell muliply both within and at the edges of the colony as shown by time-lapse cinematography. Upon long-term cultivation, *nu/nu* mice. Fetal and adult hepatocyte lines derived from day -2 fetuses were studied in detail and shown to maintain the expression of liver-specific genes for several years (Table II) cells became increasingly independent of intercellular contacts and gained the ability to multiply as individual cells, i.e. the cells became clonable. This stage correlated with the appearance of aneuploid cells. Eventually, upon further cultivation, such cells gained the ability to multiply in soft agar and displayed tumorigenic potential upon s.c. injection into nu/nu mice, leading to the development of differentiated, trabecular HCC.

Some of the steps involved in mediating increasing growth autonomy of the cells are summarized in Table III. Genetic events that underly the processes have as yet

not been characterized. They may include the production of autocrine growth factors, mutations of tumor suppressor genes and the activation of cellular oncogenes.

TABLE II

Liver-specific functions expressed in fetal mouse FMH-202 hepatocyte line after continuous culture for 34 months

Tyrosine amino transferase;	EGF receptor;
Serine dehydratase;	bFGF receptor;
Phospho-*enol*-pyruvate decarboxylase;	TGF-alpha;
Aldolase B;	IGF I;
Metallothionein	
	Liver-enriched
Ornithine transcarbamylase;	*Transcription factors:*
Argine biosynthesis	HNF1
	HNF2
Albumin	HNF3
Alpha-Fetoprotein	HNF4
Transferrin	LAP
	DAP

TABLE III

Properties of FMH-202 hepatocytes during long-term propagation in culture.

Passage number	Growth requirements	Properties
1-10	EGF & Ins (10 ug/ml) (IGF I: 10 ng/ml); Cell-cell contacts.	Diploid
10-15	Insulin (10 ug/ml) (IGF I: 10 ng/ml) Cell-cell contacts.	Diploid
20-45	Insulin (100 ng/ml) Cell-cell contacts.	Diploid
50	Abrogation of Cell-Cell Contacts	Hyperdiploid
55	Autonomy of GF &	Aneuploid of cell-cell contacts.
55-60	Growth autonomy	Growth in soft agar
>60	Growth autonomy	Tumorigenic

b) Use of transgenic FMH-202 hepatocytes for short-term mutagenicity assay systems.

Most established cell mammalian lines in culture display a limited potential to convert indirectly acting mutagens into DNA-reactive intermediates. Therefore, the metabolic activation of numerous potential xenobiotics in immortal cell culture systems widely used as short-term mutangenicity assays is commonly achieved either by co-cultivation with freshly isolated hepatocytes or by the addition of mammalian liver homogenates (e.g. S9 rat liver extracts). The transgenic FMH-202 hepatocyte line derived from day -2 fetuses was employed to explore its spectrum of metabolic activity required to generate genotoxic metabolites of indirectly acting mutagens including aflatoxin B1 (AFB1), cyclophosphamide (CP), benzo(a)pyrene

(BaP), and dimethylbenzanthracene (DMBA). The induction of sister chromatid exchanges (SCE) and clastogenic activity were examined as end points of the assays. The results of the study (Table IV) showed that FMH-202 hepatocytes efficiently activated all tested compounds as shown by clastogenic responses and by SCE induction. It was concluded, that the cells lend themselves for the identification of indirectly acting mutagens, which are converted into genotoxic intermediates by endogenous drug metabolizing enzymes in FMH-202 hepatocytes and that they can be employed for routine short-term mutagenicity assay systems *in vitro*.

TABLE IV
Clastogenic activity and induction of SCE in FMH-202 mouse hepatocytes and in V79 cells by indirectly acting mutagens

Cell type	*Compound dose M*	*CA (per sell)*	*SCE (incidence)*
V79	DMSO + S9	0.04	9.7
V79	DMBA (10−7)	0.16	15.5
FMH-202	DMSO	0.01	0.3
FMH-202	DMBA (10−7)	0.10	0.6
FMH-202	BaP (10—5)	0.13	0.6
FMH-202	CP (4 X 10−3)	0.01	0.6
FMH-202	AFB1 (10—6)	0.03	0.5

c) Use of hepatocyte line FMH-202 for transformation assays

We have begun to employ FMH-202 hepatocytes for the development of transformation assays in culture. The question was asked whether factors known to induce liver cancer in man induce untransformed
FMH-202 cells to become tumorigenic. It is well-known that persistent infection by hepatitis B virus (HBV) is epidemiologically closely associated with the prevalence of HCC in man. Transgenic mice bearing HBV DNA as a transgene that express in the liver HBSAg, develop HCC. Therefore, HBV DNA was used in our studies to probe its presumed oncogenic properties when expressed in hepatocytes. Untransformed FMH-202 cells were transfected with HBV DNA and the cultures screened for the appearance of transformed foci. Foci developed on the monolayers, which were picked and used to derive cell lines from them. All foci consisted of stably transfected cells that could be cloned. The cells displayed malignant growth characteristics in soft agar and were tumorigenic upon inoculation into nu/nu mice. The tumors were characterized histologically as anaplastic HCCs. These results provided the first demonstration that HBV displays oncogenic potential, showing that the system could be useful to functionally identify HBV genes that convey the tumorigenic phenotype to hepatocytes.

It is interesting to note that human normal hepatocytes infected in vitro by HBV could be maintained in primary culture for two months. They were able to support active viral replication all along the culture time. In spite of this permanent infection we did not find any proliferative advantage of these infected hepatocytes. Viral integration did not occur and we did not detect changes in the oxygen radical balance. These results suggested that HBV by itself was not sufficient to alter the functional activity of infected hepatocytes and does presumably not display direct oncogenic potential. This led us to postulate that some genetic alterations that

accompany hepatocyte immortalization might be required for rendering HBV tumorigenic. FMH-202 cells are useful as an assay system to identify such genetic alterations as well as HBV related carcinogenic factors.

d) Growth control of normal hepatocytes and FMH-202 immortal hepatocytes.

It is generally assumed that the G1 phase progression and the G1/S transition result from a series of transduction signals elcited in response to extracellular mitogenic factors that differ from one cell type to another, giving rise to the extreme variability of rythms of proliferation inherent to different cell types in an organism. The mechanisms that control these differences are not well understood. Early cellular changes that accompany the immortalization process are not known. The hepatocytes, which are arrested in a quiescent stage (G0) in normal liver in vivo retain the ability to divide, e.g. during regeneration of the liver following partial hepatectomy. As a model for hepatocyte proliferation, the FMH-202 hepatocyte line has been useful for analyzing the sequential alterations occurring in response to growth induction and the transformation processes.

We have analyzed both the expression of cdk1 and cdk2 throughout the cell cyle of normal proliferating rat hepatocytes. While cdk2 was constantly expressed during the cell cycle, cdk1 remained unexpressed until 20 hrs after partial hepatectomy, a time point corresponding to mid-G1, and thereafter accumulated in S, G2 and M phases. No histone H1 kinase was detected during the G1 phase, while two peaks were observed during S and M phases. These reported features defined a specific pattern of G1-associated proteins during the cell cycle of normal hepatocytes. Interestingly, we showed that the G1-associated proteins CDK1, CDK2, CDK4 and cyclins share similar patterns of expression in immortal FMH-202 cells as in normal cells. The main differences concerned the duration of G1, which appeared to be greatly reduced in the immortal hepatocytes.

In addition, we have investigated the growth factor dependence of the hepatocyte entry and progression through G1, by using primary hepatocyte cultures. It was shown that the G0/G1 transition took place during collagenase perfusion independently of mitogenic factors, wand was associated with the expression of early G1 proto-oncogenes. Evidence was then provided that normal rat hepatocytes progressed through G1 regardless of growth factor stimulation and proceeded until a restriction point located in G1. Thereafter, in the absence of mitogens, the cells remained blocked at this point and did not replicate DNA. In the presence of mitogen, hepatocytes entered S phase. However, their progression through S and M depended of the timing of an EGF signal.

In parallel, a comparative study was performed with FMH-202 cells synchronized by treatment with Na-butyrate and subsequently stimulated by growth factors. It appeared that the restriction point defined in mid-G1 in normal cells in the absence of EGF, did not exist in immortalized cells. Additional work is required to further characterize the system to identify alterations in different restriction points in G1 that accompany the development of cell types which have abrogated normal requirements for growth
factors and for cell-cell contacts, and which characterize immortalized or transformed cells.

e) Immortal transgenic cell lines from the lung

The approach to establish differentiated immortal epithelial cell lines from appropriate transgenic mice bearing an immortalizing transgene was extended by

targeting the expression of SV40 T-ag into a subpopulation of lung cells, i.e. Clara cells. The lung is known to be a highly heterogenous organ consisting of a complex mixture of about 40 different cell types, of which most are in close contact with one another. Therefore, it has as yet not been possible to establish stable cell lines of specific cell types of the lung by classical approaches.

We have employed the uteroglobin promotor to direct transcription of SV40 T-ag to the Clara cells. For that purpose, transgenic mice were established that bear gene construct UT7, in which SV40 T-ag is under control of the rabbit uteroglobin promoter. Bronchiolar cells were cultured from transgenic mouse strain UT7.1, of which permanently growing cell lines were derived. The phenotype of the resulting cell lines was as expected for Clara cells, i.e. they expressed the Clara cell specific protein (CCSP) (= uteroglobin) and surfactant protein SP-A. The mouse strain UT7.1 develops lung adenocarcinomas at the age of 3 months, the sequential growth and development of which has been studied in considerable detail. Thus, in addition to be a useful source of immortal Clara cells, the transgenic mouse strain UT7.1 represents an attractive model for human lung adenocarcinoma.

MAJOR SCIENTIFIC BREAKTHROUGHS

Transgenic mouse strains were produced that lend themselves as sources for the establishment of appropriate immortal differentiated cell lines from liver & lung, and presumably from other tissues in which the transgene is expressed.

MAJOR COOPERATIVE LINKS

Dr. Christiane Pourcell (Pasteur Inst, Paris, F); Dr. Michael Strauss (MDC, Berlin-Buch, D); Dr. Peter Jarck (H.-Pette-Institut, Hamburg, D); Dr. Gunther Schmidt (Cambes Ltd., Cambridge, GB); Dr. Michael Bauchinger (GSF, Munich, D); Dr. Lothar Hennighausen (NIH, Bethesda, Md., USA).

PUBLICATIONS

Individual publications

O'Sullivan, J.F., Schmidt, G.H., & Paul,D. Spontaneous mutation rates in the mouse. *Nature* **352**: 200-201 (1991).

Kulka, U., Paul, D., & Bauchinger, M. Development of short-term mutagfenicity test systems *in vitro*: metabolic activation of indirectly acting mutagens by theree immortal hepatocyte cell lines. *Mutagenesis* **8**: 193-197 (1993).

Gripon, P., Diot, C., & Guguen-Guillouzo, C. Reproducible high level of cultured adult human hepatocytes by hepatitis B virus: effect of polyethylene glycol on absorption and penetration. *Virology* **192**: 534-540 (1993).

Paul, D., & Halter, R. Immortal hepatocyte lines derived fromtransgenic mice expressing an immortalizing oncogene in the liver and their ude for genotoxicity and hepatotoxicity assessments *in vitro*. In: *European Medicine Research* (Ed.: G.N. Fracchia), pp. 127-139, IOS Press, (1994).

Sandmöller, A., Halter, R., Gómez-La-Hoz, E., Sulke, H.J., Paul, & Beato, M. The uteroglobin promoter targets expression of the SV40 T-.ag tio a variety of secretory epithgelial cells in transgenic mice. *Oncogene* (in press).

Sandmöller, A., Halter, R., Paul, D., & Beato, M. A transgenic mouse model for lung adenocarcinoma. *Cell Growth & Diff.* (in press).

Le Jossic, C., Glaise, D., Diot, C., Dezier, J.F., Fautrel, A., Corcos, L., & Guguen-Guillouzo, C. Trans-acting factors, detoxification enzymes and hepatitis B virus replication in a novel st of human hepatoma cell lines. (Submitted for publication).

Development of immortalized human articular cartilage cell lines for use in physiological, pharmacological and toxicological *in vitro* investigations (BIOT CT-900196)

COORDINATOR:
E.M. VEYS, G. VERBRUGGEN, Dept. Rheum., Univ Hosp Ghent, B
PARTICIPANTS:
M. ADOLPHE, Laboratoire de Pharmacologie cellulaire, EPHE Paris, F
G.S. PANAYI, UMDS, Rheum. Dis. Unit, Univ. London, London, GB
L. DE RIDDER, Histology Dept, Univ. Ghent, Ghent, B
A. CLOAREC, Département de Biologie, Laboratoires UPSA, F

BACKGROUND INFORMATION

Connective tissue cells, e.g. chondrocytes (CHO), are responsible for repair of connective tissues in pathological conditions. Their activity is affected by various biological mediators and pharmacological substances. CHO may also secrete the autoantigens that are involved in the pathogenesis of rheumatoid disorders. *In vitro* investigations on these cells should provide better understanding of the physio(patho)logy of connective tissue. However, human articular CHO are difficult to obtain and isolated cells have to be propagated to obtain sufficient numbers of cells for experimental studies. Passage numbers of nonimmortalized cells are limited and dedifferentiation (altered phenotype) of the propagated cells during culture remains a pitfall. Immortalization of these cells solves the problem of poor availability. Immortalized human CHO can be infinitely propagated. Adapted culture methods for these cells have to be developed.

OBJECTIVES AND PRIMARY APPROACHES

To investigate the possible use of isolated human CHO as a target cell in *in vitro* systems for the study of tissue breakdown and repair in inflammatory conditions. To develop immortalized human CHO lines which retain some of their differentiated functions as target cells in these *in vitro* test systems. To study the phenotypical stability of these immortalized CHO. To characterize CHO autoantigens in rheumatoid arthritis and to prepare a CHO expression library from human chondrocytes.

RESULTS AND DISCUSSION

A. Culture of human articular CHO as phenotypically stable cells

Articular cartilage was obtained from cadaveric human knee joints. CHO were isolated after enzymatical digestion of the extracellular matrix. Culture of cartilage cells in gelified matrices (agarose, collagen, alginate) prevented these cells from dedifferentiating to fibroblasts. CHO suspended in agarose were shown to remain phenotypically stable and to synthesize cartilage-specific extracellular matrix macromolecules: type II collagen and aggrecans. More than 90% of the polysaccharide material synthesized by articular CHO in agarose were sulfated aggrecans. Chondroitin-sulfate was the main glycosaminoglycan in these aggrecans. Small amounts of hyaluronan were produced, which were used as backbone molecules in the aggrecan aggregates. Electron microscopic studies confirmed the

presence of aggrecans and collagen fibers in the immediate surroundings of the cells.

Aggrecan aggregates were released from the CHO cultures and visualized after rotary shadowing. Mature human articular CHO predominantly synthesized small-size aggregates with 3 to 12 aggrecans per aggregate. However, a few large aggregates containing more than 100 aggrecans per aggregate were observed in each culture.

***Metabolism of human CHO* in vitro**: As in *in vivo* conditions, synthesis rates of aggrecans were inversely correlated with the age of the donor. These synthesis rates can be down- or upregulated by omitting or adding different growth factors from or to the culture medium. Replacing serum by bovine serum albumin reduced aggrecan synthesis to 20-30% of the basal synthesis rates. Synthesis levels were restored by supplementing the serum-free media with insulin at 100 ng/ml and with TGF-beta at 10 ng/ml.

Both pro-inflammatory cytokines, Il-1 and TNF-alpha, significantly reduced aggrecan synthesis by the cells and induced the secretion of neutral proteinase and collagenase, thereby provoking extracellular matrix degradation. CHO *in vitro* thus reacted in a predictable manner and behaved as in the *in vivo* situation. The effects of various exogenous (e.g. pharmacological) agents were investigated in this model.

B. The use of isolated chondrocytes in *in vitro* models for the study of tissue breakdown and repair in inflammatory condition

Peripheral blood monocytes (PBM) were kept in monolayer culture in pyrogen-free conditions and remained in a nonactivated state. On exposure to bacterial LPS these PBM were activated to release quantifiable levels of IL-1, TNF and of other inflammatory products in the supernatant media. Articular CHO were cultured in the presence of these activated PBM. At this time, aggrecan synthesis by the cartilage cells became significantly reduced and increased proportions of low molecular weight aggrecans were detected in the incubation media. Catabolism was induced by the secretion of proteinase and collagenase by the Il-1/TNF-triggered cartilage cell. The latter was also responsible for the release of huge amounts of Il-6 in the incubation media. This coculture model allowed -1- to identify and to accurately quantify different noxious products liberated by inflammatory cells, -2- to measure the degradation of the different parts of the extracellular environment and -3- to study the involvement of the connective tissue cell in degradation as well as in defence mechanisms.

The effects of the drugs and of the various growth factors and cytokines on extracellular matrix metabolism were reproduced with cartilage cells from different donors. This particular culture system for phenotypically stable chondrocytes and coculture of these cells with inflammatory cells seemed particularly suited to study the effects of pharmacological substances and of biological mediators in inflammatory rheumatic conditions.

C. Immortalization of human articular cartilage cells

The polyomavirus t° sensitive (ts) A mutant large T of SV40 was used to immortalize the cartilage cells. The cells were cotransfected with the pSV2tkNeo$_b$ plasmid, which encodes for the resistance of the transfected cell to G418 (geneticin). The major difficulty was the latency or crisis period the cells entered in after transfection. This period lasted for six to nine months. After this time some cells began to proliferate and clones of immortalized human chondrocytes

were established. CHO from 9 donors were transfected. Clones were obtained from 3 of these donors.

1) SV40 large T antigen was detected by immunofluorescence. For some clones, 100% of the cells were intensively labelled. For others, labelling intensity varied according to the fields observed. Southern blot analysis allowed to verify the integration of the immortalizing oncogene in the cellular genome and showed that the number of copies integrated was very different from one clone to another.

2) Loss of specific function during propagation: Collagen synthesis rates were approx.10 to 20 times lower than those of nonimmortalized monolayer-cultured human CHO. Indirect immunofluorescence staining was used to detect type II and I collagen and showed a wide variety of phenotypes. Purified ^{3}H-collagen (alpha) chains were studied after 6% SDS-PAGE electrophoresis. SDS-PAGE disclosed the presence of collagen I, I'trimer and II but did not allow to differentiate between these types. Most of the clones secreted type III collagen. It was concluded that most immortalized CHO-lines had lost their specific function when propagated in monolayer culture.

3) Assessment of differentiated function of immortalized CHO was done in culture systems that allowed the culture of phenotypically stable cells (culture in gelified matrices, e.g. agarose). Proteoglycan synthesis was studied after exposure of the cells to ^{35}Sulfate. Proteoglycan synthesis levels were reduced to approx. 25% of the values of age-matched nonimmortalized chondrocytes.

Gel chromatography of the ^{35}S-proteoglycans showed a small proportion of aggregated macromolecules. Biochemical analysis of the proteoglycan glycosaminoglycan chains showed 50% of this material to be dermatansulfate. The glycosaminoglycans were poorly sulfated. The CHO-type of proteoglycan was thus replaced by a fibroblast type of proteoglycan. It can be concluded that the cells had dedifferentiated to fibroblasts.

During long-term culture of the cells in gelified agarose (up to 12 weeks) the synthesis rates increased significantly. The cells still produced dermatansulfate-containing proteoglycans (fibroblast products) and the CHO-phenotype did not improve.

However, when Northern blot analysis of type II collagen mRNA expression was done, the signal enhanced after the cells had been cultured for 3 weeks in suspension culture in another artificial matrix (alginate beads).

4) **In vitro** ***study of extracellular matrix metabolism by immortalized CHO in an inflammatory condition*** **in vitro**: IL-1 depressed proteoglycan synthesis and induced catabolism by immortalized CHO in agarose. Cocultured activated PBM only weakly affected extracellular matrix metabolism.

D. Chondrocyte (CHO) autoantigens in the etiopathogenesis of rheumatoid arthritis (RA)

RA is a chronic inflammatory joint disease of unknown etiology, in which pannus formation results in the degradation of articular cartilage and bone. Following total knee joint replacement, in which the articular surfaces are removed, synovitis disappears in the operated but persists in the nonoperated joints of the same patient, suggesting that an antigen responsible for the etiopathogenesis of RA has been 'surgically' removed. The nature of these antigens is unknown. CHO maintain the structure and function of cartilage by regulating the synthesis and degradation of

its components. Recent evidence suggests that CHO express autoantigen(s) that may be implicated in the etiopathogenesis of RA.

Isolation and characterization of chondrocyte autoantigens: A human chondrosarcoma HTB94 cell line and human articular cartilage libraries were screened with immunoaffinity purified sera from patients with RA, osteoarthritis (OA) and relapsing polychondritis (RP). Protein extracts prepared from HTB94 cells were separated on SDS polyacrylamide gels and electro-blotted onto nitro-cellulose membranes. An enhanced chemiluminescence detection system was used to detect antibodies directed against CHO-specific antigens found in the above-mentioned sera. The isolation and characterization of autoantigens detected by Western blotting was carried out using an expression library of HTB94. A Uni-Zap human chondrosarcoma HTB94 expression library was constructed. The integrity of the expression library was assessed by screening with a beta-actin cDNA probe and monoclonal antibody and subsequently with RA sera. The positive recombinants were excised from the Uni-Zap XR vector by the ExAssist method, which allowed direct sequencing of the genes of interest directly from the Bluescript Phagemids. The DNA sequence(s) of these putative CHO autoantigens was/were compared to GENBANK/EMBL databases.

Three serum autoantibodies to the human chondrosarcoma cell line, HTB94, were identified in patients with RA. Over 60% of RA sera contained IgG antibodies which recognized a 65 kDa CHO protein. The identity of this CHO antigen is not known. The antigen had no reactivity against fibroblast extracts and it was definitely not the mycobacterial 65 kDa heat shock protein, HSP65. Several other autoantibodies directed against CHO proteins, but not against type II collagen, were detected in the serum of RA and OA patients.

A different pattern of bands was observed with OA sera, in which 20% of sera detected two proteins of 96 kDa and 80 kDa. This method is now being applied to detect autoantibodies in the serum of RP patients and obviously has potential for application to other human and even animal arthritides.

To characterize these cartilage autoantigens a Uni-Zap XR expression library has been constructed from mRNA isolated from HTB94. This library has been screened with RA and immunoaffinity purified RA sera. Three recombinants have been isolated. These positive clones have been identified by nucleotide sequence analysis. Comparison of the nucleotide sequence of two of these clones to the EMBL/GENEBANK databases has shown them to be unique genes. Consequently, they could be RA-CHO specific autoantigens. The third recombinant is to some degree similar to a T cell clone.

MAJOR SCIENTIFIC BREAKTHROUGHS

— *In vitro* systems for the study of tissue breakdown and repair
The system seems particularly suited for the study of the effects of pharmaceuticals and of biological mediators in inflammatory rheumatic conditions.

— *In vitro* systems for the study of tissue breakdown and repair in inflammatory conditions
These systems could replace some experimental animal arthritis models.

— Immortalization of human articular cartilage cells
The use of immortalized CHO as an alternative for non-immortalized CHO for the study of connective tissue metabolism in *in vitro* models is still under investigation.

Chondrocyte (cho) autoantigens in the aetiopathogenesis of rheumatoid arthritis
The definition of chondrocyte autoantigens in the pathogenesis of human rheumatic disorders is a major development in our understanding of these diseases and could lead to disease-specific immunotherapeutic intervention.
In the search for CHO autoantigens 3 CHO-specific proteins were shown to react with rheumatoid arthritis (RA) sera, while 2 other proteins reacted with osteoarthritis (OA) sera. These proteins will be used to study T-cell responses of patients with RA, OA and control subjects.

A number of chondrocyte-specific proteins may be newly discovered autoantigens. The definition of chondrocyte autoantigens in the pathogenesis of human rheumatic disorders is a major step forward in our understanding of these diseases and could lead to disease-specific immunotherapeutic intervention.

The treatment of rheumatic disorders absorbs a substantial part of the health care budget. There are no *in vitro* systems that give sufficient information on the etiopathogenesis of cartilage destruction and the effects of pharmaceutical substances on this process. Well-defined connective tissue cell lines, i.e. cartilage cells, suited for studies in the field of degenerative and inflammatory joint diseases, are now made available. Efforts are being made to redifferentiate these chondrocytes, so that their function in *in vitro* models mimics their function *in vivo*. Coculture systems with these chondrocytes as target cells and cells that participate in inflammation have been prepared. These systems could replace some experimental animal arthritis models.

MAJOR COOPERATIVE LINKS

Human chondrocytes supplied by the Ghent Rheumatology Department (A) have been immortalized in the Paris Laboratory for Cellular Pharmacology (B) and dispatched to the other centers for studies on proteoglycan metabolism (A) and collagen metabolism (B), for electron microscopic evaluation (Ghent Histology Department) and for the construction of ex expression library (London Rheumatic Disease Unit).
Chondrocytes from cartilage of laboratory animals with experimental joint diseases (UPSA Lab, Paris) have been obtained by A and B for metabolic investigations.

PUBLICATIONS

Joint publications

M. Cornelissen, G. Verbruggen, A.M. Malfait, E.M. Veys, M. Dewulf, P. Hellebuyck, L. De Ridder: Size distribution of native aggrecan aggregates of human articular chondrocytes in agarose. *In Vitro* Cell Dev. Biol. 29A, 356-358, 1993.
M. Cornelissen, G. Verbruggen, A.M. Malfait, E.M. Veys, C. Broddelez, L. De Ridder: The study of representative native aggrecan aggregates synthesized by human articular chondrocytes *in vitro*. J. Tiss. Cult. Meth. 15, 139-145, 1993.
J.B. Benoit, A. Tachet des Combes, S. Thenet, S. Demignot, M. Adolphe, A.M. Malfait, G. Verbruggen, E.M. Veys, M. Dewulf: Immortalisation de chondrocytes articulaires humains: caractérisation de différents clones et étude de leur phénotype. 3e congrès de la S.P.T.C., Paris, 5-6 March 1992.
G. Verbruggen, B. Benoit, A.M. Malfait, A. Tachet des Combes, S. Thenet, S. Demignot, M. Adolphe, E.M. Veys: Development of an immortalized human articular cartilage cell line. Phenotypic characterization of several clones. Arthritis Rheum. 9 (suppl) S318, 1992.
G. Verbruggen, A.M. Malfait, K.F. Almquist, E.M. Veys, S. Thenet, B. Benoit, S. Demignot, A. Tachet des Combes, M. Adolphe: Development of immortalized human articular cartilage cell lines. In: Joint destruction in arthritis and osteoarthritis. Eds. Van den Berg et al., Birkhäuser Verlag Basel, AAS 39, 1993, 267-272.

B. Benoit, A. Tachet des Combes, G. Verbruggen, S. Demignot, A.M. Malfait, S. Thenet, E.M. Veys, M. Adolphe: Immortalization of human chondrocytes by temperature sensitive SV40 large T antigen. Phenotypic characterization of several clones. Osteo arthritis and Cartilage, 1, 1-2, 1993.
A.M. Malfait, G. Verbruggen, E.M. Veys, J. Lambert, L. De Ridder, M. Cornelissen: Comparative and combined effects of interleukin 6, interleukin 1 beta, and tumor necrosis factor alfa on proteoglycan metabolism of human articular chondrocytes cultured in agarose. J. Rheumatol. **21**, 314-320, 1994.

Individual publications

P. Jobanputra, V. Corrigall, G.H. Kingsley, G.S. Panayi: Cellular responses to human chondrocytes: absence of allogeneic responses in the presence of HLA-DR and ICAM-1. Clin. Exp. Immunol. **90**, 336-344, 1992.
P. Jobanputra, V. Corrigall, G.S. Panayi: Expression of the 65 kD heat shock protein in human chondrocytes. Brit. J. Rheumatol. **32**, 259-260, 1992.
G. Verbruggen, A.M. Malfait, E.M. Veys, L. Gyselbrecht, J. Lambert, K.F. Almqvist: Influence of Interferon-gamma on isolated chondrocytes from human articular cartilage. Dose dependent inhibition of cell proliferation and proteoglycan synthesis. J. Rheumatol. **20**, 1020-1026, 1993.
G. Verbruggen, A.M. Malfait, M. Dewulf, C. Broddelez, E.M. Veys: Standardization of nutrient media for isolated human articular chondrocytes in gelified agarose suspension culture. Osteoarthritis and cartilage, 1994, (in press).
A.M. Malfait, G. Verbruggen, K.F. Almqvist, C. Broddelez, E.M. Veys: Coculture of human articular chondrocytes with peripheral blood mononuclear cells as a model to study cytokine-mediated interactions between inflammatory cells and target cells in the rheumatoid joint. *In Vitro* Cell Dev. Biol., 1994, (in press).
A.M. Malfait, G. Verbruggen, K.F. Almquist, J. Lambert, E.M. Veys: A model for the co-culture of human articular chondrocytes with peripheral blood mononuclear cells. Osteoarthritis and cartilage, 1, 8, 1993.

The risk potential of biotechnology-derived polypeptides causing nephrotoxicity based on assessment in human and animal renal cell lines using molecular and cellular biology techniques (BIOT CT-910266)

COORDINATOR:
P. BACH, University of East London, London, GB
PARTICIPANTS:
H. STOLTE, Nephrologie, Medizinische Hochschule Hannover, Hannover, D
J.-P. MORIN, INSERM, Uer Medicine Pharmacie, St-Etienne du Rouvray, F
M. RYAN, Dpt Pharmacology, University College, Dublin, IRL
European Free Trade Association Laboratories
J. DIEZI, Institut de Pharmacologie de l'Université de Lausanne, Lausanne, CH
W. PFALLER, Institute of Physiology, University of Innsbruck, Innsbruck, A
SUBCONTRACTORS:
CIT, Evreux, F
H. LENTZEN, AG MADAUS, Köln, D
M. SOOSE, University of Giessen, D

BACKGROUND INFORMATION

Renal cell cultures offer a unique system to test and understand mechanisms of nephrotoxic xenobiotics. Freshly isolated fragments or primary cultures have a limited lifespan and lose their *in vivo* characteristics. However, continuous or permanent cell lines maintain certain functions, but lack specific metabolic, transport and receptor characteristics which could be important to toxicity screening.

OBJECTIVES AND PRIMARY APPROACHES

The major goals were to

- characterize and improve the state of differentiation of available renal epithelial cell lines by development of new culture conditions,
- tailor new, non-transformed and non-tumorigenic cell lines by fusion of partner-cells expressing different functional properties,

and elaborate *in vitro* test procedures based on these cell lines to effectively and sensitively detect deterioration of function or injury.

RESULTS AND DISCUSSION

A. Optimised phenotypic expression of rabbit proximal tubule cell primary cultures

Rabbit proximal tubule cells are generally grown in the presence of glucose, insulin and a bicarbonate buffer containing phenol red. The omission of insulin and glucose prevents glycolysis and delays decreased gluconeogenesis and cells have a high density of microvilli on the apical surface as is found *in vivo*. Replacing the bicarbonate buffer by 10 mM Hepes buffer pH 7.4 restores the mitochondrial succinate dehydrogenase activity to levels comparable with fresh proximal tubules. The uptake of phenol red into proximal cells alters phase II detoxication.

Thus reoptimised culture conditions shows biochemical and functional improvements through the modulation of culture composition and the environment

that make the rabbit proximal tubule primaries more phenotypic of these cells *in vivo*.

1. Comparative characteristics of cell lines of 'proximal tubule origin'

The LLC-PK_1 (pig kidney), LLC-RK_1(rabbit kidney) and OK (opossum kidney) cells lines, grown either in the presence of FCS or in hormonally defined media were compared to primary rabbit proximal tubules. Based on a range of marker enzymes representative of key biochemical processes, optimised rabbit primary cells exhibited the closest phenotype to the *in vivo* situation for cells that were assessed.

2. Impact of phenotype expression on the toxicity response

Optimised primary cultures of rabbit proximal tubule cells represented the closest *in vitro* model to the *in vivo* situation for the study of aminoglycoside-, cephalosporine- and cisplatin-induced nephrotoxicity than LLC-PK_1 and OK cells.

3. Tailoring new cells

A new OK cell line, capable of performing gluconeogenesis and the transport of organic anions has been derived. Fusion of cell lines (rabbit proximal tubule cells and LLC-PK_1 and OK cell lines failed to provide hybridoma cell lines expressing the sum of the properties of their ancestor cells and phenotypic improvements, and fusion of hypoxanthine guanidine phosphoribosyltransferase negative LLC-PK_1 cell lines with primary human proximal tubular cells are still being functionally and morphologically characterized.

4. Assessment of cellular function

A number of novel techniques have been used to assess renal cell function.

5. Epithelial barrier function.

The electrophysiological measurement of paracellular ion conductivity is a sensitive way of assessing loss of epithelial barrier function. Interferon alpha has been shown to have a direct effect on epithelial cells and gentamicin toxicity was prevented by magnesium aspartate hydrochloride. Measurement of areas of 'domes' represents unidirectional ion and solute transport in solid support grown renal epithelial cell monolayers. This is also a sensitive indicator of deterioration of epithelial barrier function caused by toxic xenobiotics or peptidic drugs. The most sensitive cellular system appears to be the lysosomal compartment releasing activities of N-acetyl-β-D-glucosaminidase in parallel to changes in the paracellular conductivity of epithelial monolayers.

6. Flow cytometric techniques

Flow cytometry offers rapid and accurate measurement of cellular viability, cell fusion, alterations in intracellular calcium and pH, cell cycle and DNA profiles and apoptosis using cell size and appearance of a 'sub-G1' peak. In general, LLC-PK_1 cells (proximal tubular origin) are more sensitive than MDCK cells (distal tubule origin) to the nephrotoxicity of gentamicin and CsA, where increased intracellular Ca^{2+} precedes loss of cellular viability and provides a common event in many forms of toxic damage. Low doses of CsA produced a 'sub-G1' peak and cell shrinkage in LLC-PK_1 cells, suggesting apoptosis.

7. Fluorescent probes

Heavy metals such as Hg^{2+} and Ag^{+} (as low as 10^{-7}M) induced a dramatic increase of transepithelial Na current within seconds, probably due to the induc-

tion of 'channels' in the apical membrane, which have low ionic selectivity. The adding sulphydryl donors protect against this effect, suggesting the involvement of -SH groups. Cd^{2+} ions induce similar, but less efficient and potent effect that was unchanged by -SH groups. These metals also increased intracellular Ca^{2+} either rapidly for Hg^{2+} or slowly for Cd^{2+}. The effects for Ag^{+} was delayed and oscillations markedly. This signal was very sensitive to chelators.

A range of cells have been characterised in terms of the expression of biochemical characteristics or functions in the presence or absence of nephrotoxic chemicals. This approach has allowed the early detection of oxidative effects, mitochondrial injury, changes in the Golgi apparatus and the shift of expression of protein kinase C between the nuclei and cytoplasm.

B. Glomerular toxicity

Disorders of glomerular barrier function occur in renal injury and diseases. These have been associated with an impaired metabolism of the mesangial matrix constituent fibronectin.

1. Isolated glomeruli

Adriamycin (ADR) affects the multicellular isolated glomeruli of the Atlantic hagfish (*Myxine glutinosa*) by inhibiting protein synthesis and/or proteolytic enzymes, which in the absence of increased RNA-synthesis suggests inhibited transcription and translation.

Rodent glomeruli show similar sensitivity to ADR, and the use of fluorescent probes and biochemical methods demonstrate that redox cycling and oxidative stress are not likely to be key to the degenerative processes that take place in these cells.

2. Mesangial cell cultures

ADR alters the turnover of fibronectin, a mesangial extracellular matrix protein which has been studied in human mesangial cell cultures. ADR increases steady-state fibronectin mRNA levels and decreased degradation of fibronectin released from the cells into the culture media. As a consequence of proteinase inhibition fibronectin accumulates extracelluarly, which provides negative feedback control of translational expression.

Rodent mesangial cells in culture (and a range of other cell types) show changes in the expression of protein kinase C which is most effectively inhibited by ADR. If mesangial cell signalling is significantly changed for any length of time it is likely that this could contribute to the progressive degenerative injury associated with glomerular disease.

MAJOR SCIENTIFIC BREAKTHROUGHS

Establishment of

— culture conditions ensuring optimum expression of *in vivo* phenotypes in established renal cells and hybridoma lines.

  immortal gluconeogenic OK line and several yet uncharacterized hybridoma lines, obtained by fusing cell lines and primary proximal tubular cells.

— electrophysiologic methods to assess disturbances of epithelial barrier function by measuring transepithelial ion conductivity in renal epithelial cell monolayers grown on microporous supports,

— quantitative morphological methods to assess epithelial barrier function in monolayers grown on solid supports,

— cellular markers (enzymes and proteins) released into the culture medium that show cell injury in monolayer cultures,

— the use of fluorescent probes for flow cytometry, ion visualisation and the assessment of organelle and cytoplasm function and biochemistry that helps us to understand the mechanisms of cell injury.

Pre-validation step and/or industrial applications

— *Electrophysiological assessment* has been used in several laboratories and could be developed as a screening test for detecting nephrotoxic compounds. As a result of the work carried out in this contract the test is ready for transfer to prevalidation and possible validation.

— *Flow cytometry* facilitates molecular mechanistic toxicity studies in addition to assessment of gross cytotoxicity. The work during this contract indicates that flow cytometry can be used in screening and assessment of toxic compounds and a novel putative mechanism of cyclosporin A toxicity involving apoptosis was detected.

— *Fluorescent probes* have facilitated the identification of organelle and cytoplasmic changes within cells exposed to toxic chemicals and the modulation of kinase signalling systems.

— *Multicellular systems*. Whole glomeruli and their cultured component cells have allowed a better understanding of causal relationships of ADR cytotoxicity and of interspecies differences in response to the drug.

MAJOR COOPERATIVE LINKS

These are too extensive for a brief report, but have involved each of the collaborating laboratories, where expertise, personnel and materials have been exchanged.

The group has also set up collaborative links with a range of academic and state funded research institutions and commercial organisation involved in drug development and the sale of equipment and consumables for cell culture in West, Central and Eastern Europe and North America.

PUBLICATIONS

Joint publications

Fillastre J.P., et al., Pheochromocytoma in a hemodialysis patient. *Am. J. Kidney Dis.*, **19**: 94-95, 1992.

Gstraunthaler G., et al., Ammoniagenesis in LLC-PK_1 cultures: role of transamination. *Am. J. Physiol.*, **263**: C47-C54, 1992.

Gstraunthaler G, et al., (1990) Renal cell cultures: a tool for studying tubular function and nephrotoxicity. *Toxicology Letters* **53**, 1-9

Gstraunthler, G. et al., A novel gluconeogenic strain of OK cells with metabolic properties different from gluconeogenic LLC-PK_1 cells. *Cell Physiol. Biochem.*, **3**: 78-88, 1993.

Kastner, S., et al., Metabolic heterogeneity of isolated cortical and juxtamedullary glomeruli in adriamycin nephrosis. *Renal Physiol. Biochem.* **14**: 48-54, 1991.

Kastner, S., et al., Metabolic studies on isolated rat glomerular; a valuable tool to investigate glomerular damage. In: Nephrotoxicity: Mechanism, Early diagnosis and Therapeutic Management. (Bach, P.H., et al., eds.) Marcel Dekker, New York, p. 467-473, 1991.

Soose, M., et al., A comparative study on the response of two established cell lines to adriamycin. Growth and biomass production in human mesangial cell cultures and in Tetrahymena thermophile. *Eur. J. Protistol.*, **30**: 75-84, 1994.

Soose, M., et al., Fibronectin turnover in human mesangial cell cultures as affected by adriamycin. *Cell Biol. Toxicol.* **9**: 149-164, 1993.

Soose, M., et al., Altered renal fibronectin excretion in early adriamycin nephrosis of rats. *J. Pharmacol. Exp. Ther.* **257**: 493-499, 1991.
Soose, M., et al., Heterogeneity of glomerular barrier function in early adriamycin nephrosis of MWF rats. *Clin. Physiol. Biochem.* **6**: 310-315, 1988.
Soose, M: Properties of silibinin and of antioxidants against adriamycin cytotoxicity in an unicellular eukaryote, Tetrahymena thermophia. *Eur. J. Protistol.*, (in press).
Soose M., et al., Fibronectin turnover in human mesangial cell cultures as affected by adriamycin. *Cell Biol. Toxicol.*, (in press).
Toutain H. and Morin J.P. Renal proximal tubule cell cultures for studying drug-induced nephrotoxicity and modulation of phenotype expression by medium components. *Ren. Fail.* **14**: 371-383, 1992.
Toutain H., et al., 1991 Biochemical, functional and morphological characterization of a primary culture of rabbit proximal tubule cells. *Exp. Cell Res.* **194**, 9-18.

Individual publications

Bach, P.H. and Wilks, M.L. *In vitro* techniques for screening for nephrotoxicity, mechanism of injury and new discovery. In: In Vitro Toxicology: An Alternative to Animal Testing?, (Eds. Jolles G. and Cordier A.), pp. 59-91. Academic Press, London. 1992.
Bach P.H., et al., Fluorescent probes for the assessment of the subcellular effects of chemicals and disease processes *in vivo* and *in vitro*. In: Biotechnology Applications of Microinjection, Microscopic Imaging and Fluorescence. [Edit., Bach P.H., Reynolds C.H., Clark J.M., Mottley J. and Poole P.L.], pp. 79-98. Plenum Press, London. 1993.
Bach P.H., et al., [Editors] Biotechnology Applications of Microinjection, Microscopic Imaging and Fluorescence, Plenum Press, London. [Proceedings of the First European Workshop on Biotechnology Applications of Microinjection, Microscopic Imaging and Fluorescence, held in London in 1992]. 1993.
Barbey, M-M., et al., Adriamycin affects glomerular renal function: evidence for the involvement of oxygen radicals, *Free Rad. Res. Comms.* **7**: 195-2-3, 1989
Fels, L.M., et al., Filtration characteristics of the single isolated perfused glomerulus of Myxine Glutinosa. *Renal Physiol. Biochem.* **16**: 276-284, 1993.
Fels, L.M., et al., Free radical formation and adriamycin nephrotoxicity studies in the hagfish (Myxine glutinosa). *Bull. Mt. Des. Isl. Biol. Lab.* **28**: 22-24, 1989.
Gstraunthaler G, et al., (1993) Ammoniagenesis in renal cell culture. *Renal Physiol. Biochem.* **16**, 203-211
Gstraunthaler G, and Pfaller W (1992) Metabolic properties of gluconeogenic substrains of renal epithelial cell lines. *JASN* **3**, 828
Gstraunthaler G. and Pfaller W. Continuous renal cell lines as *in vitro* tools to study nephrotoxicity. *In vitro* Methods of Toxicology (R.R. Watson, ed.) CRC Press pp. 93-113, 1992.
Gstraunthaler G. and Pfaller W. The use of cultured renal epithelial cells *in vitro* assessment of xenobiotic-induced nephrotoxicity. In: Renal Disposition and Nephrotoxicity of Xenobiotics (Dekant et al. eds.) Academic Press pp. 27-61, 1993.
Haberstroh, U., et al., Micropuncture studies on mechanisms of proteinuria in early adriamycin nephrosis of MWF rats. in: Nephrotoxicity, Extrapolation from *in vitro* to *in vivo* and from Animals to Man. (Bach, P.H, Lock, E.A., eds.) Plenum Press, New York, p.755-760, 1989.
Kastner S., et al., Mechanistic studies on the influence of adriamycin on amino acid uptake and amino acid incorporation into isolated glomeruli of the Atlantic hagfish *Myxine glutinosa. Bull. Mount Desert Isl. Bio Lab.* **31**: 35-36, 1992.
Kastner S., et al., Nephrotoxixcity of the glomerulus. Studies with adriamycin on isolated glomeruli of the atlantic hagfish *Myxine glutinosa. Bull. Mount Desert Isl. Bio Lab.* **32**: (in press).
Kastner, S., et al., Isolated glomeruli of the Atlantic hagfish Myxine glutinosa as an alternative *in vitro* model to study glomerular metabolism in pharmacotoxicology of anticancer drugs. *Comp. Biochem. Physiol.*, (in press).
Kastner, S., et al., Human kidney cells in *in vitro* pharmaco-toxicology. In: Human Cells *in vitro* Pharmaco-toxicology. Present Status within Europe. (Rogiers, V., et al., eds.), VUBPress, Brussels,p. 197-238, 1993.
Kersting U., et al., Differentiation of Madin-Darby canine kidney cells depends on cell culture conditions. *Cell Physiol. Biochem.*, **3**: 42-55, 1993.

Kersting U, et al., (1993) Differentiation of Madin-Darby Canine Kidney Cells Depends on Cell Culture Conditions. *Cell Physiol Biochem* **3**, 42-55
Leroy A., et al., Penetration of lomefloxacin into human prostatic tissue. *Am. J. Med.*, **92** (suppl. 4A): 12S-14S, 1992.
McGlynn H., et al., (1993). Flow cytometric assessment of gentamicin nephrotoxicity in established renal cell lines. In: Biotechnology Applications of Microinjection, Microscopic imaging and Fluorescence. [Eds., Bach P.H., et al.,], pp. 127-130. New York Plenum Press.
Malandrin B., et al., Ccit brutale chez un malade trait par hmodialyse. *Annales de Medcine Interne*, **143**: 351-352, 1992.
McGlynn H., et al., Electrophysiological and flow cytometric studies on mechanism of gentamicin nephrotoxicity in cultured renal cells. *Arch. Toxicology. Suppl.*, **15**: 140-143, 1992.
McGlynn, H. and Ryan, M.P. Mechanisms and models of gentamicin nephrotoxicity. *Toxicology In Vitro*, 7: 407-410, 1993.
McGlynn H. and Ryan M.P. Mechanisms and Models of Gentamicin Nephrotoxicity. *Toxicology In Vitro*, (in press).
McGuinness, S.J. and Ryan, M.P. (1994). Mechanism of cisplatin nephrotoxicity in rat renal proximal tubule suspensions. *Toxicology In Vitro*, (in press).
Monteil C., *et al.*, 1993 Modulation of glycolysis induction in primary cultures of rabbit kidney proximal tubule cells: the role of shaking, glucose and insulin. *Cell Biol. Int.* **17**, 953-960.
Monteil C., et al., Insulin and glucose impact on functional and morphological differentiation of rabbit proximal tubule cells in primary culture. *Renal Physiol. Biochem.*, **16**, 212-221.
Monteil C., et al., Characterization of gentamicin induced dysfunctions *in vitro*: the use of optimized primary cultures of rabbit proximal tubule cells. *Ren. Fail.*, (in press).
Morin J.P., et al., Renal epithelial cells in culture: a comparative study of the functional activities of rabbit kidney proximal tubule cells in primary cultures and the LLC-PK_1 and OK cells. In: Animal cell technology developments processes and products. R.E. Spier, J.B. Griffiths, and C. Mac Donald eds. Butterworth-Heinemann Oxford UK — pp 713-715, 1992.
Morin J.P. et al., 1991 Rabbit kidney proximal tubule cells in primary culture: evaluation of the impact of expressed phenotype on cellular toxic response. Toxic *In Vitro* **5**, 383-388.
Moulin B., et al., Glomerulonephrotic in chronic lymphocytic leukemic and related B cell lymphomas. *Kidney Int.*, **42**: 115-127, 1992.
Netzer A, et al., (1993) Establishment of new renal proximal tubular cell lines obtained by cell fusion. *Pflügers Arch.*, **422**, Suppl 1, R91
Pfaller W., et al., Low pH and hyperosmolality determine the differentiation of MDCK cells typical for the outer medullary collecting duct. *Cell Physiol. Biochem.*, **3**: 65-67, 1993.
Pfaller W, and Willinger CC (1994) Animal and cell culture models for the assessment of acute renal dysfunction and injury. In: Clinical Nephrotoxins (Bennett WM, et al., eds.) Kluwer (in press)
Wilks M.L. and Bach P.H. Renal enzymes as the basis for alternative methods of safety screening *in vitro* and understanding the mechanism of nephrotoxicity. In: Urinary Enzymes in Clinical and Experimental Medicine, (Edits. Jung, K., et al.), pp. 299-317. Springer-Verlag, Heidelberg, 1992.
Pfaller W, et al., (1991) Hypertonicity determines differentiation of MDCK cells typical for medullary collecting duct. *Pflügers Arch.*, **418**, R69
Poirier J.M., et al., Propafenone pharmacokinetics in uremic patients treated by peritoneal dialysis. *Clin. Nephrol.*, **38**: 231-232, 1992.
Radermacher J., et al., Endothelin-3 modulates glomerular filtration rate in the isolated perfused rat kidney. *Renal Physiol. Biochem.*, **15**: 325-33, 1992.
Ryan, M.P. (1993). Interrelationships of magnesium and potassium homeostasis. *Miner. Electrolyte Metab.*, **19**: 290-295.
Steinmaäl D, et al., (1993) LLC-PK_1 cells as a model for *in vitro* assessment of proximal tubular nephrotoxicity. *In vitro* (submitted)
Thorndike J.M., et al., Chlorpromazine Cytotoxicity in 3T3-Cells. *Marine Environ. Res.*, **34**: 345-350, 1992.

AREA D:

PRE-NORMATIVE RESEARCH

BIOSAFETY

Analysis of gene transfer between microorganisms and plants (BIOT CT-910282)

COORDINATOR:
Ch.J. THOMPSON, Institut Pasteur (IP), Paris, F
PARTICIPANTS:
R. GUERRERO, Universidad de Barcelona (UBA), Barcelona, E
A. PÜHLER, Universität Bielefeld (UBI), Bielefeld, D
P. SIMONET, Université de Lyon I (UCB), Villeurbanne, F
A. TSIFTSOGLOU, Aristotle University of Thessaloniki (UT), GR
M. LEBRUN, Rhône-Poulence Agrochimie (RPA), Lyon, F

BACKGROUND INFORMATION

The purposeful release of genetically engineered organisms (GEOs) into the environment has led to increased interest in interactions between resident organisms, in particular the potential of newly introduced genetic material to transfer to new species and to become established and expressed in indigenous populations. Rhizospheres include bacterial populations classified as Actinomycetes (*Streptomyces* and *Mycobacterium*), *Rhizobium*, and *Agrobacterium* as well as a majority of microbes which cannot be cultivated in the laboratory. We are interested in finding and characterizing mechanisms by which bacteria might transfer genetic material among themselves or to receive genes from plants. Classical microbial genetic studies carried out under laboratory conditions has identified a number of transfer mechanisms e.g. transduction, transposition, conjugation, fusion and transformation, whereby genetic characters can be transferred and established in closely related recipient species. 'Trans-kingdom' gene transfer from bacteria to plants can also be demonstrated in the laboratory. However there is presently no documented proof of transfer from plants to bacteria. Nevertheless, since soil bacteria are capable of transferring genes to other microorganisms, and, in the case of *Agrobacterium*, even to plants, gene transfer from plants to soil bacteria must be seriously considered as a potential link for transfer of genetic information in the environment. It is not known whether (or at what frequency) these phenomenon actually occur in the environment. These studies were carried in the environment, in plant tumors, on conventional bacterial growth media, and in laboratory microcosms designed to mimic and allow experimental manipulation of the physiochemical characteristics of the soil.

OBJECTIVES AND PRIMARY APPROACHES

I. Assessment of possible gene transfer between plants and microorganisms under natural conditions:
   A. In a tobacco plant tumor
   B. Field studies

II. Physiochemical factors affecting conjugation efficiency

III. Bacterial conjugative mechanisms
   A. Conjugation between soil actinomycetes (*Streptomyces*)
   B. Plasmid-mediated transfer from genetically modified *Agrobacterium tumefaciens* into *Rhizobium leguminosaurum* and other bacteria

RESULTS AND DISCUSSION

I. Assessment of possible gene transfer between plant and microorganisms under natural conditions

A. In a tobacco plant tumor (UBI)

We chose the soil bacterium *Agrobacterium tumefaciens* as recipient for gene transfer for the following reasons:

— The tumor induced by agrobacteria is a natural habitat which brings *A. tumefaciens* and plants cells into intimate contact.

— The transfer of DNA from *A. tumefaciens* to the plant is well known and analyzed. (Nevertheless, to date there is no evidence that the transfer system functions only in one direction.)

— Agrobacteria carry sequences homologous to those transferred to the plant via *Agrobacterium*-mediated gene transfer (e.g. T-DNA-borders). This might facilitate the establishment of transgene-carrying DNA fragments from the plant in the *Agrobacterium* by homologous recombination.

Transgenic plants were constructed which carried the *aacCI* gene conferring gentamicin (Gm) resistance in bacteria, the *npt*II coding region conferring kanamycin resistance in plants and the luciferase coding region (*luc* from *Photinus pyralis*) which is expressed in plants and bacteria. *Agrobacterium* wild-type strain A281 (Gm^s/*luc*$^-$) was chosen as recipient.

In our first experiments, involving $> 10^{12}$ bacteria, we attempted to transfer the Gm^r/*luc* genes to A281 using either transgenic plant tissues or their chromosomal DNA. Despite numerous experiments in which A281 was incubated with purified transgenic plant DNA (in the presence and absence of various plant extracts) or with leaf discs or extracts, Gm^r, *luc*$^+$ transformants were not observed.

Analysis of 2×10^{11} bacteria isolated from 76 tumors containing the *aacCI* and *luc* genes generated approximately 4000 Gm^r clones all of which proved to be *luc*$^-$. Southern hybridization analysis of 22 of these isolates showed that they did not contain the *aacCI* gene. In addition, Gm^r strains could be isolated from tumors of wild type plants not containing the *aacCI* gene. We believe that these isolates were spontaneously resistant to Gm rather than transgenic organisms. Hence, under these conditions which closely simulate natural bacteria-plant interactions, inter-kingdom gene transfer was not observed.

B. Field studies (RPA)

Characterization of transgenic plants and approval for field trials. Transgenic tobacco plants were constructed by *A. tumefaciens* (LBA 4404)-mediated transformation. The T-DNA contained a selectable marker (the kanamycin resistance gene under the control of the *nos* promoter and terminator) and one of two reporter genes with bacterial regulatory signals: the gentamicin resistance gene *aacC1* (525-type plants) or bromoxynil resistance gene *oxy* (523-type plants).

Western analysis of proteins extracted from these two lines using antibodies raised against these proteins did not show a positive signal (analysis of the 525-4 plants were done by the IP group). Moreover, the 523-3 plant was killed after treatment with low doses of bromoxynil. Although the prokaryotic promoters did not allow expression in these tobacco plants, recloning the *aacC1* gene from the plant in pCU19 conferred gentamicin resistance in *E. coli*. This demonstrated that the gene

was still functional upon transfer to a bacterial host. Similar experiments with the 523-3 plants were inconclusive since we were unable to detect expression of the *oxy* gene in *E. coli* (ability to use the bromoxynil as the sole nitrogen source for growth).

Authorization for release of the 525-4 plants (*aacC1*-containing) in field trials [(Bergerac (SEITA) and Chazay d'Azergues (RPAgro)] was obtained from the French Commission du Génie Biomoléculaire. Experiments were carried out in order to detect gene transfer during aging of tobacco leaves and to provide a defined source of soil sample for future comparative studies.

1. Development of detection techniques. Sensitive methods have to be developed for detecting specific genetically defined microorganisms within complex natural microbial communities and for monitoring the fate of their DNA in natural ecosystems. However, the complexity and the heterogeneity of most soils hampered the development of such technologies, resulting in detection tests lacking sensitivity. We developed a new method based on the polymerase chain reaction (PCR) for the detection of bacteria in the soil. In order to improve the sensitivity of the detection we have optimized each step of the protocol including lysis of cells, DNA purification and PCR amplification. The microorganisms in the soil were successively treated by sonication, micro-wave heating, and the thermal shocks to maximize the efficiency of the lysis. Purification of nucleic acids was achieved by passage through up to three Elutip-d columns. Finally, PCR amplifications were optimized via biphasic protocols using booster conditions, lower denaturation temperatures and in the presence of formamide. This method has been developed on 100 mg soil samples inoculated with *A. tumefaciens* suspensions. Specific primers were characterized in the plasmid-borne *vir* genes and have routinely permitted the detection of 10^7 to 10^3 inoculated *A. tumefaciens* cells. Moreover, the strong correlation we observed between the size of the inocula and the yields of the PCR reactions permitted assessment of the validity of the protocol in enumerating microbial cells inoculated into a soil sample.

However, living and dead cells, as well as extracellular DNA released into the soil after the death of the microorganisms are expected to provide targets for the PCR reaction. In order to restrict the analysis to DNA from microbial cells, a technique has been developed to specifically extract bacteria from the soil matrix (without isolating the bacteria on culture media) and to purify it from extracellular DNA contaminating the surface of these cells. Intracellular DNA corresponding approximately to 10^7 bacteria per g of soil could then be extracted, amplified and enumerated.

2. Persistence of DNA in soil may favor gene transfer. Adhesion of microbes and macromolecules to soil components is known to be related to surface characteristics of both partners, such as charge, hydrophobicity, and microaggregate formation, and to pH and soil water content. These factors are also involved in the adhesion of nucleic acids from plant, animal or microbial cells, which are released into the soil after their death.

The understanding of mechanisms involved in DNA adsorption onto soil particles could provide information for monitoring the fate of free DNA in the soil, including its ability to be transferred to soil microbes via transformation. We investigated the relationships which could exist between clay montmorillonite with high surface areas and the DNA. The extent of DNA adsorption was affected by the concentration and valence of the cations used (Mg^{2+}, Ca^{2+}, Na^{+}), indicating a charge dependent process. Calf thymus DNA was found to be highly adsorbed by smectite

(up to 30 mg/g of dry clay). Adsorbed DNA was shown to be more resistant to degradation by DNase I than free DNA. Experimental data with plasmid and short linear PCR-amplified DNA showed that protection against nucleases was only partial. Nevertheless, clay-adsorbed DNA was found to be stable, even after a strong DNase I treatment, to artificially transform competent *E. coli* cells.

In order to study the persistence of the *aacC1* gene from transgenic tobacco plants in soil, and its potential transfer to the indigenous microflora, a series of protocols were developed based either on the extraction of total DNA from soil or from bacteria isolated from soil. Soil samples were collected in the field one year after harvesting the plants. Selective primers have been characterized to specifically amplify DNA from:

Plants — using the SSU RuBisCO region
Bacteria — using a ribosomal DNA gene
Trangenic plants — using the reporter gene *aacC1*.

Our PCR results showed that tobacco plant DNA could be detected in the soil confirming the ability of soil components to protect extracellular DNA. Moreover, the use of primers complementary to part of the *aacC1* gene also allowed amplification of template DNA extracted from soil. However, the question remains open as to whether this DNA originated from bacteria naturally harboring this gene, from contaminant plant material or from indigenous microorganisms transformed with the recombinant plant DNA.

We also examined transformation of naturally competent *P. stutzeri* cell with homologous DNA in soil microcosms ranging from pure clay minerals to natural soil samples. Natural transformation was found to occur in most of the experiments including the soil extract microcosms. These results confirm that even in a complex medium such as soil, adsorbed and protected DNA can transform naturally competent bacteria.

II. Physiochemical factors affecting conjugation efficiency (UBA)

We were interested in defining ecological factors which may influence conjugal transfer with the ultimate aim of using the information to predict the likelihood of gene transfer at sites where GEOs could be released. Conjugation was studied in *Rhizobium* species and *E. coli* using the microcosms/metal-pin sampler procedure developed during the previous Biotechnology Action Program (BAP) project. Because of its simplicity and reproducibility, our tests in microcosms also generated accurate data describing the mobility of a given population.

The physiological state of bacterial cells is probable one key factor in determining transfer frequency. Poly-β-hydroxyalkanoate (PHA), an energy and carbon polymeric reserve compound in most bacteria, has proven to be an good indicator of the physiological status of microbial communities in nature. Optimal environmental conditions for growth are associated with low PHA content, whereas adverse environmental conditions for growth are associated with high levels of PHA. Experiments in soil microcosms showed that PHA-rich *R. meliloti* cells have slightly higher conjugation frequencies than PHA-poor cells.

Our results showed that conjugation frequency was also dependent on:

— *soil water content* — Both *E. coli* and *R. meliloti* conjugation was most efficient insoil having a water content of about 20%. *E. coli* conjugation efficiency was more sensitive to water content than *R. meliloti*. This was probably related to

previousdata showing that survival and dispersal of *E. coli* cells in soil were low comparedwith *R. meliloti.*

— *temperature — The optimal temperature was around 30°C.*

— *pH- Rhizobium meliloti* dispersion rates in soil were altered only at highly acidic (pH < 5) or alkaline conditions (pH > 9), whereas *E. coli* strains showed a stenotopic response. However, gene transfer frequencies in both species were nearly constant (conjugation was slightly most efficient in the neutral pH range).

III. Bacterial conjugative mechanisms (IP, UT)

Previous BAP (IP) studies demonstrated that Gram-negative bacteria (*E. coli*) can transfer genetic material to Gram-positive bacteria (*Streptomyces* and *Mycobacterium*). These RP4 and RSF1010 conjugative processes rely on some specific transfer functions supplied in trans as well as an origin of transfer carried by the conjugative plasmid.

A. Conjugation between Streptomyces (IP)

Streptomyces are Gram-positive filamentous eubacteria representing a major population in soil ecosystems and as such they could be involved in environmental gene flux. Fertility (gene exchange based on homologous recombination) in *Streptomyces* is unusual in that it is dependent on plasmid transfer and both parents apparently function simultaneously as donor and recipient of chromosomal genes. At the outset of these studies, we believed that this effect may not involve transfer of the chromosome; plasmid genes may only trigger chromosomal recombination after hyphal fusion has brought the content of the two parental cytoplasms into a single compartment. The formation of such heterokaryons was suggested by genetic studies carried out in the 50's and 60's. In this study, we tried, using tools of modern molecular biology to corroborate the existence of heterokaryons in *Streptomyces* mixed populations and to identify the potential plasmid gene(s) which may control recombination.

Electron micrographs of mixtures of *Streptomyces lividans* and *E. coli* or *Mycobacterium smegmatis* and *E. coli* under conditions where conjugation is taking place have revealed bridges between the two organisms. The structure of these points of contact suggest a fusion-like event which may allow mixing of the two cytoplasms or even genetic recombination between the two organisms. However, these points of contact between Gram-positive actinomycetes and Gram-negative *E. coli* are not dependent on plasmid-specific transfer functions.

In order to determine whether hyphal fusion results in cytoplasmic mixing, we subcloned the genes encoding the *Vibrio harveyi* luciferase subunit a or b into two different vectors. As expected, when these genes were cloned together on different plasmids in the same host, light was produced. Similarly, we thought that heterokaryon formation would generate light. However, using this highly sensitive and specific reporter system we did not detected any photons indicating mixing of cytoplasms under conditions favoring 'heterokaryon formation'.

In order to determine whether a specific plasmid gene promotes chromosomal recombination, we subcloned the a PCR-synthesized *tra* gene known to be essential for intra-mycelial transfer under control of a thiostrepton inducible promoter (*ptipA*) so that the expression of *tra* was controlled by the presence (*tra* turned on) or absence (*tra* turned off) of thiostrepton. After many unsuccessful attempts, one plasmid, was isolated which allowed thiostrepton-inducible plasmid transfer. However nucleotide sequence analysis of the cloned *tra* gene revealed that the it

had undergone gross rearrangement. We suppose that the *tra* gene product is extremely toxic and that even low level of expression of *ptipA*, in non induced cultures, is enough to prevent recovery of viable strains carrying the desired clone.

B. Plasmid mediated transfer from genetically modified A. tumefaciens *into* Rhizobium leguminosaurum *and other bacteria (UT)*

We developed methods to assess possible transfer of RP4 derivatives carrying the *aacC1* gene from genetically modified *Rhizobium leguminosaurum* into *A. tumefaciencs*, and *Pseudomonas syrignae* in our previous BAP studies. This work were extended here to investigate whether RP4 derivatives could be transferred from genetically engineered *A. tumefaciens* (UBAPF2a,b) into *R. leguminosaurum* natural isolates collected from soil samples in Greece and into other bacteria. Based on the genetic and growth characteristics of the *R. leguminosaurum* isolates (GNI), a few of them were selected as potential recipients for conjugation with *A. tumefaciens*. (UBAPF2a,b). In most, if not all, of the conjugation experiments carried out with various *R. leguminosaurum* GNIs as recipients, no transconjugants were obtained. However, when UBAPF2b *Agrobacterium* was crossed with *R. leguminosaurum* GNI3d or GNI1d, candidate transconjugants were observed. Rigorous genetic testing of these potential transconjugants revealed that these were not bona fida transconjugants.

We also investigated whether RP4 derivatives could be transferred from *A. tumefaciencs* UBAPF2b into *E. coli* XL-1 and into *Staphylococcus aureus* strains. Conjugation experiments between UBAPF2b and *E. coli* and/or *S. aureus* indicated that while no mating occurred between *A. tumefaciencs* and *S. aureus*, potential transconjugants were obtained from crosses of UBAPF2b and *E. coli*. Assessment of the transconjugants for growth in fortimicin, expression of *aacC1* enzyme activity, expression of resistance to antibiotics, presence of plasmid DNA, restriction mapping of bacterial genome with endonucleases as well as Hofle RNA genotyping indicated transfer into *E. coli* cells under laboratory conditions. These conjugation experiments have shown that although RP4-mediated transfer of *aacC1* from genetically engineered *A. tumefaciencs* into *R. leguminosaurum* GNIs and Gram positive bacteria was not possible, low frequency ($1x10^{-6}$) transfer occurred into *E. coli*. Although the mechanism(s) of this exchange of RP4 derivatives between GEOs and other bacteria are not known, the data show that genetic exchange between GEOs and other bacteria can occur at low frequency.

MAJOR SCIENTIFIC BREAKTHROUGHS

DNA of trangenic tobacco plants can be detected by PCR in fields at least one year after the plants had been harvested. No evidence for gene transfer from plants to bacteria was obtained.

MAJOR COOPERATIVE LINKS

Perhaps the most productive link was established between Rhône Poulenc Agrochimie and the Université Claude Bernard in their field studies which will be continued in the future. The University of Bielefeld has traditionally served as a center providing genetic advice and strains to other members of the group.

PUBLICATIONS

Joint publications

Tirodimos I., Pretorius-Guth, I.-M., Preifer, U., Tsaftaris, A. & Tsiftsoglou, A.S. (1993): Methods of assessment of transfer of a gentamicin-resistance gene (*aacC1*) from genetically engineered microorganism(s) into agriculturally important soil bacteria. Appl. Microbiol. Biotechnol. 38:256-530.

Paget, E., Jouan, D., Lebrun, M., and Simonet, P. (1994): Monitoring the fate of recombinant plant DNA in soil. Appl. Environ. Microbiol. (in preparation).

Individual publications

Broer, I., Dröge-Laser, W. & Gerke, M.: Estimation of the potential horizontal gene transfer between plants and agrobacteria. (In preparation).

Gascon, J.Oubina, A., Pérez-Lezaun, A. & Urmeneta, J. (1994): Sensitivity of selected bacterial species to UV radiation. (Submitted).

Guerrero, R. & Mas-Castella, J. (1994): Gene transfer in the environment. In: Microbiology and Molecular Genetics. J. Casadesus (editor). Spanish Society for Microbiology. (In press).

Kyriakides, X., Gennimata, D., Tsaftaris, A. & Tsiftsoglou, A.S.: Assessment of transfer of a gentamicin acetyltransferase gene (*aacC1*) from genetically engineered *Agrobacterium tumefaciens* into *E. coli* and other bacteria. (In preparation).

Lafuente, R. (1994): Genetic transfer between microorganisms in soil microcosms: Effects of environmental conditions. M. Sci. Thesis. Univ. of Barcelona, Spain.

Lafuente, R., Urmeneta, J. & Mas-Castella, J.: Gene transfer between GEMs in microcosms. 14th Spanish Congress on Microbiology. Zaragoza (Spain). September (1993).

Lafuente, R., Maymo, X., Mas-Castella, J. & Guerrero, R. (1994): Influence of environmental factors on plasmid transfer in a soil microcosms. (Submitted).

Mazodier, P. & Thompson, C. J. Plasmid transfer and fertility functions of pIJ101 (in preparation).

Paget, E. & Simonet, P. (1994): On the track of natural transformation in soil. FEMS Microbiol. Ecol. (In press).

Paget, E., Jocteur-Monrozier, L. & Simonet, P. (1992): Adsorption of DNA on clay minerals: Protection against DNAse I and influence on gene transfer. FEMS Microbiol. Lett., 97, 31-40.

Picard, C., Ponsonnet, C., Paget, E., Nesme, X. & Simonet, P. (1992): Detection of bacteria in soil by direct DNA extraction and PCR. Appl. Environ. Microbiol. 58, 2717-2722.

Pretorius-Güth, I. & Pühler, A. (1993): The established alfalfa nodule does not facilitate horizontal gene transfer from *Rhizobium meliloti* to the host genome. Microb. Releases 2, 143-151.

Fate of genetically engineered micro-organisms (GEMs) and genetically engineered DNA sequences (GEDs) in some environmental hot spots (BIOT CT-910284)

COORDINATOR:
A. M. MERGEAY, V.I.T.O., Mol, B (lab A)
PARTICIPANTS:
B. J. FRY, University of Wales, Cardiff, GB (lab B)
C. J. MARQUEZ, I.B.E.T., Oeiras, P (lab C)
D. W. VERSTRAETE, University of Gent, Gent, B (lab D)

BACKGROUND INFORMATION

GEMs are increasingly used in the open field for bioremediation and agricultural purposes. Knowledge of their survival in the environment and of their interactions with organisms indigenous to the environment are mandatory to predict safety and efficacy. Genetic interactions between GEMs and micro-organisms that are autochtonous to the environment are of special importance as they may allow GEMs to survive longer than predicted and to spread GEDs into microbes that are better adapted to the environment. Accordingly, GEMs may acquire genes that increase their environmental fitness (e.g. heavy-metal resistance genes in heavy-metal polluted areas) and GEMs may receive environmentally derived conjugative plasmids that may mobilise GEDs back to the indigenous microflora (e.g. by retrotransfer). This project aimed to assess the probability that such genetic interactions occur in two environmental spots, i.e. soils (in presence/absence of plants and/or pollution) and in river stone biofilms.

OBJECTIVES AND PRIMARY APPROACHES

Microcosms that simulate soil (with or without plant roots) and river stone surface environmental spots were seeded with GEMs containing marker genes (GEDs) cloned in different vector systems. Vector systems (Mob^+ and Mob^- plasmids, chromosome) differed in their capacity to be mobilised by broad-host range conjugative plasmids. The following parameters were followed:
(i) survival of GEMs and stability of their corresponding GEDs,
(ii) transfer of GEDs into co-introduced recipient bacteria,
(iii) transfer of GEDs into indigenous bacteria,
(iv) uptake of indigenous conjugative plasmids by GEMs, and
(v) effect of GEMs on the metabolic capacity of the indigenous microflora.

RESULTS AND DISCUSSION

1. Monitoring of GEDs in soil

To study the fate of GEDs in soil, a method was needed that detects GEDs irrespective whether these are present in GEMs, in indigenous microbes or in a cell-free state. A soil DNA extraction procedure was developed that allows isolation of free and cell-associated DNA of sufficient purity for hybridisation and PCR analysis [lab A]. The method consists of a lysozyme/SDS treatment of soil slurry at pH 9.0 and of a purification of the extract by ammonium acetate precipitation and Sephadex G50 gel filtration. With this method up to 30 mg of soil can be analysed in a single PCR reaction without apparent inhibition by soil contaminants. The method showed its practical value in an experiment in which soil microcosms were seeded with *Alcaligenes eutrophus* cells and exposed to high temperature or

to desiccation. The number of *A. eutrophus* colony forming units that could be recovered from the microcosm quickly declined in function of time whereas the concentration of a typical *A. eutrophus* plasmid gene remained unaltered, as determined by hybridisation analysis of soil DNA extracts (reference 3). Thus, GEDs may be more stable in soil than their corresponding GEMs [lab A].

2. Gene transfer between GEMs and co-introduced recipient cells in microcosms

Soil microcosm were seeded with both GEMs (i.e. donor) and recipient cells and the transfer of GEDs from GEMs into the recipient cells was monitored. Two donor/recipient pairs were studied into detail:

A first mating pair consisted of *E. coli* K12 donor cells with *czc* (an *A. eutrophus*-derived gene cassette that confers resistance to cadmium, cobalt and zinc) as the marker gene and *czc*-free *A. eutrophus* recipient cells. As *czc* is only functional in *A. eutrophus* but not in *E. coli*, transfer can easily be monitored by the appearance of metal-resistant bacteria. *E. coli* belongs to the γ-Proteobacteria and *A. eutrophus* to the β-Proteobacteria, thus transfer is considered to be of the 'broad-host range' type. As the vector that carries the GEDs is important for its 'transferability', *czc* gene cassettes were cloned in different vector systems, including pUC18 (Tra^-, Mob^-), pBR325 (Tra^-, Mob^-), pRK415 (Tra^-, Mob^+, IncP), pKT240 (Tra^-, Mob^+, IncQ), and the chromosome of *E. coli* (obtained by homologous recombination within the *mal*T — *mal*Q genes) [lab A]. The following are the conclusions obtained from studies with this mating pair in soil microcosms: Transfer occurs frequently if all the following conditions are met:

(i) *czc* is present on a Mob^+ vector,
(ii) tra functions are provided in the donor or in the recipient cell, and
(iii) the soil is sterilised prior to seeding with donors and recipients.

Transfer of *czc* on Mob^- vectors may occur if a strong transposable element (Mini-Mu) is present in donor or recipient. Pollution of the microcosm with heavy metals was sometimes found to induce a strong proliferation of *czc*-containing transconjugants [lab D].

A second mating pair consisted of *E. coli* K12 XL-2 donor cells (carry a F' plasmid with a tetracycline resistance gene) and *E. coli* K12 RB85 recipient cells and belongs to the 'narrow host range' type of gene transfer. Both *E. coli* strains were found to colonise the plant roots of *Zea mays* grown in sterilised soil. Gene transfer between both *E. coli* strains occurred in sterilised soil with *Zea mays* plants if the microcosm was kept at high temperature (28°C night, 32°C day). However, no gene transfer was detected when plants were omitted or when the temperature was sub optimal (18°C night, 28°C day). Switching microcosms conditions from unfavourable to favourable (plants present, high temperature) resulted in a delayed transconjugant formation. The presence of plants was not needed for transconjugant formation when soils were amended with nutrients. Gene transfer was shown to be reduced in the presence of the montmorillinitic clay bentonite. This reduction was more pronounced at low than at high ionic strength [lab C].

3. Transfer of GEDs into the indigenous microflora.

In this series of experiments, GEMs were introduced in microcosms and transfer of GEM plasmids into the resident microflora was followed. To increase the chances of success, use was made of GEMs that contain conjugative plasmids (Tra^+).

Soil microcosms were studied using *E. coli* GEMs that contained an IncP plasmid with a biphenyl degradation transposon (Tn*4371*). The biphenyl degradation gene was found to be functional in *A. eutrophus* and some *Acinetobacter* sp. strains but

not in the GEM itself. The GEM was introduced in microcosms of sandy and sandy-loam soil that were freshly amended with biphenyl. Indigenous transconjugants that had taken up the IncP plasmid were detected on basis of their capacity to grow on minimal agar plates with biphenyl as sole carbon source (donors are unable to grow on biphenyl and are methionine auxotrophs). Soil transconjugants carried plasmids that did not differ in size from those of the GEM. No transfer was observed in control soil samples that were not amended with biphenyl [lab D].

Microcosms consisting of agar plates with biofilm from River Taff stones (=epilithon) were seeded with *P. putida* PaW340 GEMs with an epilithic IncP plasmid (pQM901) that contains a mercury resistance gene (*mer*), a kanamycin resistance gene and the *tdn*C gene (encoding conversion of catechol to a yellow cleavage product). The GEMs were tryptophan auxotrophs. After co-culture of the GEMs with the epilithic micro-organisms, mercury-resistant, kanamycin-resistant tryptophan prototrophs were isolated. Isolates that carried a plasmid of the same size as pQM901, converted catechol into a yellow cleavage product, hybridised to the *tdn*C gene and had a whole cell protein finger print that was different from that of the GEMs, were scored as indigenous transconjugants. Transconjugants were identified by using API20NE and API20E test strips as *Klebsiella oxytoca, Serratia liquefaciens, Serratia plymuthica,* and *Serratia ficaria* [lab B].

River stream microcosms were also seeded with *P. putida* PaW340 (pQM901). After mating, donor counter selection was performed by incubating the donor/microcosm mixtures in presence of two different *P. putida* PaW340-specific bacteriophages. The treated mixture was then plated on LB-based medium supplemented with mercury and kanamycin. As a result, a number of indigenous transconjugants were obtained that scored positive by whole cell protein profiling, hybridisation and plasmid visualisation [lab B].

4. Transfer of conjugative plasmids from the selected environmental spots into the GEMs.

In the previous paragraphs it was shown that GEMs may transfer GEDs cloned in Tra^+ vectors into the indigenous microflora of soil and of river epilithon. However, most GEMs used for deliberate releases in the environment do not contain conjugative plasmids. Therefore transfer of GEDs into the natural microflora is unlikely unless GEMs are able to take up conjugative plasmids from the environment.

Uptake of conjugative plasmids by GEMs was studied by exogenous isolation, a technique based on the in vitro mating of bacterial communities with adequately marked recipient strains. A first approach of exogenous isolation is based on the observation that in many environmental (conjugative) plasmids mercury resistance genes are present. Mercury-sensitive, rifampicine-resistant *P. putida, P. aeruginosa, E. coli* and *Acinetobacter calcoaceticus* GEMs were co-cultured on agar plates with river epilithon bacteria and mercury-resistant GEMs were isolated. GEMs so-obtained contained either large (50-300 kb) or small (< 10 kb) plasmids. About one fourth of the plasmids isolated in *E. coli* also transferred to *P. putida*. The host range of a typical plasmid isolated by exogenous isolation into *P. putida*, i.e. pQM706, was compared with that of an IncP plasmid (pQM899). This analysis showed that both plasmids could transfer to other members of the γ-Proteobacteria group (*P. putida, A. calcoaceticus, E. coli*), and to members of the β-Proteobacteria (*Comomonas acidovorans, Pseudomonas cepacia, A. eutrophus*). However, pQM899 could transfer into the flavobacterium *Cytophaga heparina* whereas pQM706 could not. Thus, conjugative plasmids with a broad transfer spec-

trum are present in river epilithon and can be isolated by the exogenous isolation technique [lab B].

A modified exogenous isolation technique was developed that specifically detects broad-host range conjugative plasmids. For this purpose, an environmental microbial community is mixed with a donor strain that contains an IncQ (Tra$^-$ Mob$^+$) plasmid and with a recipient strain. Transfer of the IncQ plasmid into the recipient strain can only occur if the microbial community contains IncQ-mobilising functions. By choosing *E. coli* (γ-Proteobacteria) as a donor and *A. eutrophus* (β-Proteobacteria) as a recipient, only IncQ-mobilising plasmids with a broad host range are isolated. The IncQ plasmid in the donor carries the *czc* gene cassette as a marker gene. As discussed before the *czc* cassette is not functional in *E. coli* and many environmental bacteria but is in *A. eutrophus*. The 'triparental exogenous isolation method' has the additional advantage that its success does not depend on the presence of a marker gene such as *mer* on the conjugative plasmid. A control experiment showed a linear relationship between the number of IncP plasmid containing helper cells in the mating mixture and the number of transconjugants that eventually emerged. Hence, the method may be used to quantify IncQ-mobilising functions in environmental samples. With soil and activated sludge microbial communities as helper, *A. eutrophus* transconjugants were obtained that, in 50% of the cases investigated, contained a plasmid in addition to the IncQ plasmid. Further mating experiments suggested that these additional plasmids could mobilise the IncQ plasmid to other *A. eutrophus* strains and, in some cases, also to *E. coli*. Replicon typing using DNA probes specific for several *E. coli* incompatibility groups (IncN, IncP, IncQ, IncW, IncT, IncX, IncU, IncF, IncH1, IncHI2, and IncL/M) showed that all strains with two plasmids hybridised to the IncQ probe, but only 4 of 23 hybridised to the IncP probe and 0 of 23 hybridised to the other hybridisation probes. This suggested that in soil and activated sludge broad-host range IncQ-mobilising plasmids are present that belong to hitherto uncharacterized incompatibility groups [lab A and D].

5. Effect of GEMs on the metabolic capacity of the indigenous flora

Experiments were performed to analyse whether introduction of GEMs into soil and river epilithon microcosms could influence the metabolic activity of the indigenous organisms. The oxygen uptake of a biphenyl-polluted soil microcosm that had taken up the biphenyl-degradation genes from an introduced *E. coli* GEM (see point 3) was found to be higher than the oxygen uptake of unseeded soils. This suggested that GEDs transferred into the indigenous population were indeed expressed and altered the metabolic activity of the soil microbes [lab D]. In another series of experiments the capacity of unseeded and GEM-seeded river epilithon microcosms to degrade a large panel of carbon sources was compared. The GEM carried a conjugative plasmid (pQM901) that was not expected to confer selective advantage to its recipients. No significant differences in carbon utilisation patterns were found. The effect of this GEM on the growth rate of *Gammarus pulex* placed in the river microcosms was also tested. *G. pulex* is a small freshwater invertebrate that feeds on epiphytic bacteria. No significant differences between control and seeded microcosms were observed. No GEMs or indigenous microbes containing pQM901 were found in *G. pulex* gut homogenates [lab B].

6. Discussion

This project has investigated the question: 'How realistic is the possibility that genetic interactions between GEMs and environmental microbes may affect the stability of GEDs in the environment'. First, we have shown that GEDs may some-

times remain present in soil microcosms for a longer time than the corresponding GEMs (Point 1). In plate matings with soil and river epilithon micro-organisms GEMs were found to acquire conjugative plasmids from environmental microbes (Point 4). From GEMs with conjugative plasmids GEDs were found to be mobilised to the indigenous microflora (Point 2 and 3). Both gene transfer into and gene transfer from the GEMs may occur in a single mating (retrotransfer) or in two separate matings (one with an indigenous microbe that acts as donor and one with another that acts as recipient). Depending on the environmental constraints and the biological containment of the gene, such transfer of GEDs from GEMs to indigenous microbes may occur frequently (e.g. Mob^+ plasmid gene, positive selective pressure and nutrients available, few competitors) or rarely if at all (chromosomal gene, no nutrients, many competitors).

MAJOR SCIENTIFIC BREAKTHROUGHS

- A newly-developed isolation method for conjugative plasmids (exogenous isolation) showed the presence of new groups of broad-host range plasmids in soil and in the bacterial biofilm present on river stones.
- Observation that GEDs can be transferred by conjugation from introduced GEMs into the indigenous bacteria of soils and river epilithon.
- Observation that the IncP plasmid pQM899 can transfer from *P. putida* into *Cytophaga heparina*.
- Development of a DNA purification procedure suitable for the routine analysis of soil samples by PCR.

MAJOR COOPERATIVE LINKS

GEMs and GEDs were constructed in lab A and used in lab C and lab D to monitor gene transfer (results of lab C and lab D were compared). Exogenous plasmid isolation techniques originally developed in lab B were improved in lab D. Plasmids isolated in lab D were characterised in lab A. Staff of lab D worked for short periods in lab A and lab B. Semestrial meetings were held in Mol (19.03.1992), Oeiras (22.10.1992), Cardiff (23.03.1993), and Grenada (24.10.1993). Joint progress reports were presented at the Sectorial Biosafety Meeting in Wageningen (06-09.12.1992) and Grenada (24-27.10.1993). Occasional contacts (exchange of strains and protocols) took place with partners of other BRIDGE projects, including BIOT CT-910293, BIOT CT-910288 and BIOT CT-910301. The European collaborative work enriched a whole range of activities in the partners' laboratories, especially as a result of exchange of different approaches. In particular it was felt that maximal benefits appeared after long-term collaborations (>3 j).

PUBLICATIONS

Joint publications

Top E., Van Rolleghem P., Mergeay M., and Verstraete W. (1992)
Determination of the mechanism of retrotransfer by mechanistic mathematical modelling. J. Bacteriol. **174**: 5953-5960

Top E., De Smet I., Verstraete W., Dijkmans R., and Mergeay M. (1994)
Exogenous isolation of mobilizing plasmids from polluted soils and sludges. Appl. Environ. Microbiol. **60**: 831-839.

Individual publications

Dijkmans R., Jagers A., Kreps S., Collard J.M., and Mergeay M. (1993). Rapid method for purification of soil DNA for hybridization and PCR analysis, Microbiol. Releases, **2**: 29-34.

Springael D., Diels L., Hooyberghs L. Kreps S., and Mergeay M. (1993). Construction and characterization of heavy-metal-resistant haloaromatic degrading *Alcaligenes eutrophus* cells. Appl. Environ. Microbiol. **59**: 334-339.

The effects of selection on gene stability and transfer in populations of bacteria in soil (BIOT CT-910285)

COORDINATOR:
A. E.M. WELLINGTON, Dept of Bio. Sc., Univ. Warwick, Coventry, GB
PARTICIPANTS
B. A. KARAGOUNI, Univ. of Athens, Athens, GR
C. A. AKKERMANS, Dept Microbiology, Agric. Univ., Wageningen, NL

BACKGROUND INFORMATION

Antibiotic resistance is used as a marker for many bacterial inocula and in addition is found in antibiotic-producing streptomycetes grown industrially for antibiotic production. It is necessary to evaluate the effects of introducing antibiotics into soil to study the selective advantage of such genes. In addition we need to consider the safe disposal of recombinant streptomycete waste from the pharmaceutical industry.

OBJECTIVES AND PRIMARY APPROACHES

The overall objective was to study the effects of selection on introduced antibiotic resistance genes in soil populations of streptomycete inoculants by the addition of antibiotics to soil. The aim was to determine selective effects in soil microcosms using genetically marked inoculants with resistance genes borne on plasmids, phage and the chromosome. Soil microcosm design was investigated to establish a suitable way of making sporadic amendments to the soil and obtain a population of streptomycetes which followed the complete life cycle from germinating spore to mycelium in soil. The effects of soil type on survival and selection were also determined as three different soil types were chosen representing a characteristic soil of the three participating countries. Two antibiotics, neomycin, a water soluble antibiotic, and thiostrepton, a water insoluble antibiotic, were chosen for the study. Antibiotics similar to these two are frequently used in agriculture for growth promotion in farm animals and are applied to land via use of contaminated animal manure. Methods for antibiotic extraction and detection were developed to achieve sensitive assays for monitoring the rate of decay or immobilisation of antibiotics added to soil. Evidence for antibiotic production in soil was also sought using streptomycetes producing neomycin and thiostrepton. In addition methods were developed for direct monitoring of bacterial growth and differentiation in soil by differential DNA extraction and amplification of low levels of DNA by the polymerase chain reaction (PCR).
Division of labour: Microcosms were developed by each participant but continuous enrichment systems were produced and used only by Athens. Extraction and detection of thiostrepton was developed by Warwick and methods for neomycin by Wageningen. Selection and gene transfer were investigated for plasmid-borne genes at Wageningen and Athens and for chromosomal genes at Warwick. Marked inoculants were used throughout and indigenous marked strains used in conjunction with type strains for experiments at Athens. Three soil types were identified as representative of the agricultural soil within the vicinity of each participating laboratory, the soils differed markedly in terms of clay and organic matter content.

RESULTS AND DISCUSSION

All strains were tested for stability of genetic markers which proved acceptable for plasmid, phage and chromosomal insert except that studies had to be discontinued with the transposon due to instability. For soil microcosms studies with thiostrepton proved straightforward but problems were encountered with the water soluble neomycin. No antibiotic could be extracted from soil except in the case of the Wageningen soil when it was not limed. This soil was acidic and had to be limed to obtain reproducible data on growth of inoculants. Levels as high as 1000 $\mu g\ g^{-1}$ soil were not detected in the clay loam soil of Athens, in the sandy loam soil of Wageningen detection of levels below 500 μg^{-1} were not possible.

1. ***Genetic markers:*** The genes *tsr, nptII* and *aph V* for resistance to thiostrepton and neomycin respectively were used in various constructs to determine the effects of selection on resistance phenotype. Derivatives of the plasmid pIJ101, pIJ673 and pIJ680, were used for plasmid studies and KC301, derived from C31, served as the phage vector both as free virions and in lysogens. A strain of *Streptomyces lividans*, TK64(KT), carrying a chromosomal insert with *tsr* and *nptII* was chosen for studies of chromosomal genes. All studies with inoculants involved marked strains of *S. lividans*, TK23, TK24 selected by respective resistance to spectinomycin and streptomycin. Plasmids were used by Wageningen and Athens whereas phage and chromosomal genes were selected for study at Warwick. Indigenous *S. griseus* strains with rifampicin resistance were also used by Athens.

2. ***Dynamic soil microcosms:*** The selective pressures of antibiotics in the soil environments could only be determined if inoculant streptomyces were actively growing and achieved repeated rounds of sporulation. Spores were resistant to the antibiotics. This requirement led to the development of a stirred fed-batch microcosm where 20% of the soil was removed and replaced by fresh amended soil at two week intervals. This resulted in iterative germination and sporulation of streptomycetes in the soil. Whilst this system was adopted by all laboratories Athens also developed a series of microcosms which were amended or stirred for a range of time intervals. The effect was the same in all systems where either nutrient amendment or stirring resulted in spore germination. The soil conditions were held constant for moisture and temperature. One exception to this were studies completed by Athens where the effects of soil matric potential on survival and activity of plasmid-containing strains was monitored. Results showed that survival was better in Greek soil with a potential of -300 kPa when compared with soil at-23 kPa. Extraction studies proved that soil DNA from the drier soil contained higher levels of plasmid pIJ673 (Karagouni et al., 1993)

3. ***Antibiotic selection on*** *Streptomyces* species: The effects of addition of thiostrepton and neomycin to soil were studied for a range of *Streptomyces* species.

a) *The effects of thiostrepton:*
Thiostrepton was extracted from soil and detected throughout the experiment using thiostrepton-specific ultra-sensitive bioassays and further confirmed by HPLC, levels declined by 50% over 60 days but fluctuated with the soil amendments.
The three soil types all gave comparable results with thiostrepton but killing effects were most marked in Wageningen soil and least in Warwick soil.
Plasmid studies at Wageningen indicated that sensitive recipients were inhibited and no growth occurred above $50\mu g\ g^{-1}$ soil except after 30 days when the thiostrepton level had declined despite two weekly 20% soil amendments (Fig. 1). It is possible the streptomycetes were able to degrade this antibiotic as a source

of C and N. After 30 days plasmid transfer resumed but transconjugant population levels were much reduced (Fig. 1).

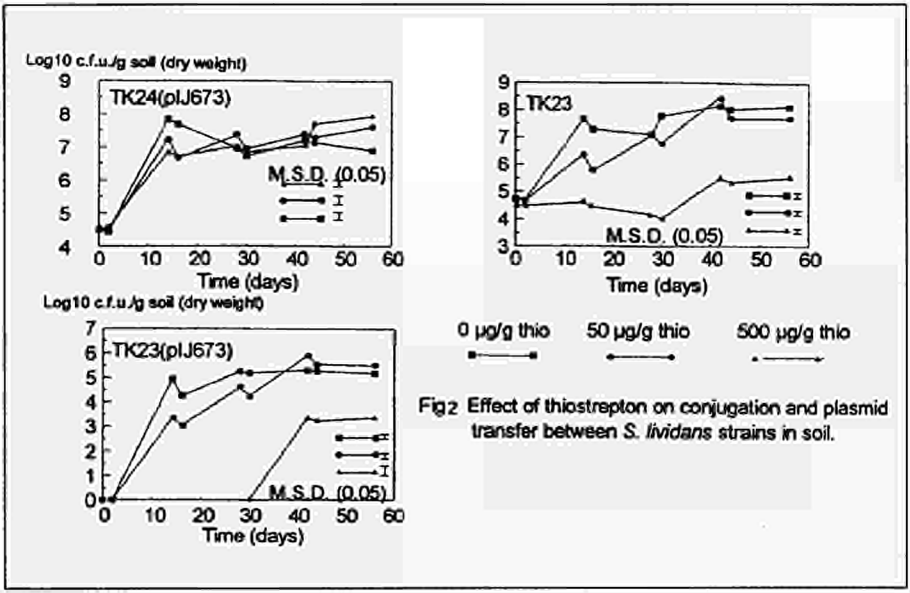


Fig2 Effect of thiostrepton on conjugation and plasmid transfer between *S. lividans* strains in soil.

0-1000µg g^{-1} thiostrepton were added to sterile batch microcosms (clay loam Warwick soil) containing the thiostrepton-sensitive *S. lividans* TK23. The antibiotic caused limited inhibition but appeared to enhance sporulation at 50, 100 and 1,000 µg g^{-1} (Fig. 2). This suggested that a response to premature death of mycelia produced from germ tubes, in this case, was increased spore formation. TK64(KT), a strain of *S. lividans* with chromosomal thiostrepton and neomycin resistance genes, was unaffected by the addition of thiostrepton.

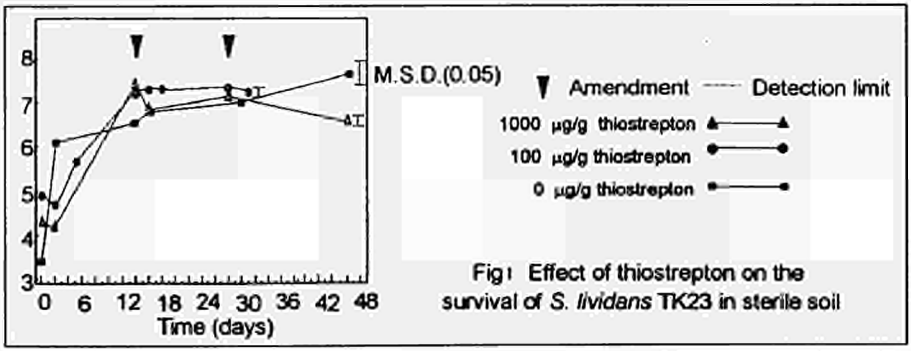
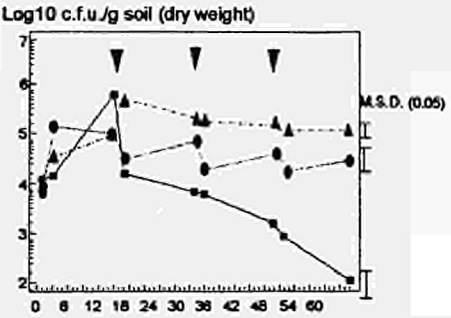


Fig1 Effect of thiostrepton on the survival of *S. lividans* TK23 in sterile soil

The strain TK64(KT) survived only in sterile soil but declined and became non-detectable after 60 days in non-sterile soil (Fig. 3). However the addition of 100μg g^{-1} or greater levels of thiostrepton dramatically altered the ability of this auxotroph to survive and compete with indigenous soil bacteria as shown in Fig. 3.



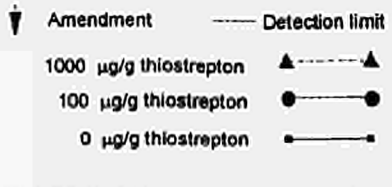


Fig 3 Effect of thiostrepton on the survival of *S. lividans* KT in non-sterile soil

The selective effect of thiostrepton on a mixed population of TK23 and TK64(KT) was investigated in sterile. TK23 was unaffected by the presence of TK64(KT), while the latter declined either in the presence or absence of thiostrepton. The initial decline in spore counts for TK23 was due to germination and evidence for killing was found at this time but the surviving mycelia appeared to sporulate more profusely and thereby aided the survival of TK23.
The survival of thiostrepton-resistant TK24, lysogenized by KC301 containing *tsr*, was not enhanced by the addition of thiostrepton and this inoculant was out-competed by TK23 in soil both with and without antibiotic additions. The results indicated that lysogens grew and/or sporulated less effectively and that although some lethal effects were monitored in the presence of thiostrepton, enhanced and more rapid sporulation in the sensitive, non-lysogenized strain counteracted this effect.

The *tsr* marker was stable in all studies but *nptII* became silent in soil populations of S.lividans when used on the chromosome but not in the plasmid pIJ673. Probing and PCR of soil DNA for these marker genes indicated that the signal for both genes was the same (Fig. 4) but for *S. lividans* KT after 30 days in soil selection on neomycin plates was not possible.

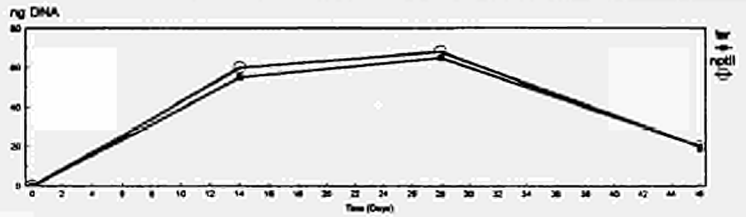


Fig 4 Quantification of genes for thiostrepton (tsr) and neomycin (nptII) resistance in S. lividans DT growing in soil

b) *In situ production of thiostrepton by S.azureus in soil:*
The growth of *S. azureus* in sterile amended (1% chitin) Warwick soil was assessed together with *in situ* production of thiostrepton. The antibiotic was detected at day 3 which coincided with a rise in spore numbers indicative of the first round of sporulation. Following 6 days incubation thiostrepton was detected at over 50 ng g^{-1} soil (Wellington et al., 1993). This level was achieved with a viable count of approximately 10^8.

c) *The effects of neomycin:* The effects of neomycin were studied with *S. griseus* CAG17 (neomycin sensitive strain) and TK24(pIJ673) using Athens soil. The microcosms in this case were stirred each day to allow continuous growth of mycelia. Under these conditions the killing effect of neomycin should be maximised as fewer resistant spores were formed. However, no evidence was obtained for lethal effects of neomycin on *S. griseus* and this strain survived in higher numbers than *S. lividans* both with and without neomycin additions (0-1000 μg g^{-1}). No transconjugants were detected. The level of neomycin in the soil could not be monitored as the antibiotic was not extracted, even at levels of 1 mg g^{-1}. Irreversible adsorption was possible as previous studies have found this occurred with several aminoglycosides. Illite can bind over 40 g of neomycin per kg and this clay is present in the soil used.

Results obtained with Wageningen soil contrasted to those reported above for the clay-loam soil. Various levels of neomycin, 0 to 1000 μg g^{-1}, were added to sterile fed-batch microcosms (sandy loam soil) containing strains TK23 and TK24 (pIJ680, aphV). In the presence of neomycin the sensitive strain TK23 fell below detection limits ($< 10^2$ c.f.u. g^{-1}) after only 30 days of incubation, while in neomycin-free soil the strain reached levels of up to 10^6 c.f.u. g^{-1}. This showed that neomycin was having a detrimental effect on the survival of TK23. TK24 (pIJ680) grew in the presence of neomycin even at 1000 μg g^{-1}. Growth of the resistant strain was greatest at the highest level of neomycin added, and this was confirmed by direct analysis of extracted soil DNA when probed for plasmid DNA. Exposure to high levels of neomycin may have affected copy number of the plasmid thus dramatically improving survival compared to the performance of the strain at lower levels of neomycin addition. The inhibitory affect of neomycin was removed when the soil was limed.

4. ***Activity measurements:*** In addition to plate counts and DNA monitoring, methods of monitoring activity have also been developed. In Warwick the production of two enzymes, urease and chitinase were used to monitor activity. Another two enzymes, agarase and amylase have also been assessed for their usefulness in monitoring populations. In Athens the respiration rate of streptomycetes in soil, measured by monitoring carbon dioxide evolution, was useful for detecting differences in activity between amended and unamended soil although the method was not sufficiently sensitive to detect the effects of stirring the soil during soil turnover.

MAJOR SCIENTIFIC BREAKTHROUGHS

Survival of a disabled thiostrepton resistant streptomycete was forced in non-sterile soil by selection for resistance using thiostrepton levels > 100μg g^{-1} soil. This has implications for industry in assessing the survival of disabled strains in soil. Particularly for mycelium contaminated with antibiotics.
Marked differences were detected between varying soil types concerning the effects of added antibiotics neomycin and thiostrepton on sensitive streptomycetes growing in soil, the presence of certain types of clay and the clay content was most

important in determining the inhibitory effects of antibiotic additions to soil. In sandy loam selection for resistant strains was observed.
The environmental implications of using antibiotic resistance genes for marking introduced organisms: the selective advantage of resistance genes in bacteria may be seen when high levels of antibiotic are present such as the levels of kanamycin in animal manure and slurries. The selective effects of such antibiotics will depend on the soil type as clay appeared to adsorb and inactivate the antibiotic neomycin.

MAJOR COOPERATIVE LINKS

Over the two years there have been four three day plenary meetings, the first in Athens 23th May 1992, the second in Wageningen 11th December 1992, the third in Warwick 17th June 1993 and finally in Wageningen 24th November 1993. A laboratory exchange was conducted in June/July 1992 between Athens and Warwick.

Exchange of skills and techniques, particularly the training of personnel in different laboratories. Opportunity to sample a wide range of soil types from differing geographical and climatic areas and use the soil in microcosm studies. Broadening our understanding of scientific approach and how this varies within Europe.

PUBLICATIONS

Joint publications

The effect of soil moisture content on spore germination, mycelium development and survival of a seeded streptomycete in soil. Karagouni A. D., A. P . Vionis, P. W. Baker and E. M. H. Wellington. Microbial Releases, 2: 47-5, 1993
Survival of streptomycetes in soil and the effects of neomycin additions on plasmid transfer. Herron, P.H., I.K. Toth, A. Vionis, A. Karagouni, A. Akkermans and E.M.H.Wellington. 1994 (in preparation).
Survival of streptomycetes in soil and the effects of thiostrepton additions on gene stability and plasmid transfer. Toth I.K., P.H. Herron, A. Vionis, A. Karagouni, A. Akkermans and E.M.H. Wellington. (in preparation).

Individual publications

The selective effects of antibiotics in soils. Wellington, E.M.H., P. Marsh, I.K. Toth, N. Cresswell, A.S. Huddleston and M. Schilhabel. In Trends in Microbial Ecology, R. Guerrero & C. Pedros-Alio, eds.,pp 331-336, Spanish Society for Microbiology, Spain, 1993
Survival of the temperate actinophage °C31 and Streptomyces lividans in soil and the effects of competition and selection on spread of lysogens. Marsh P., I. K. Toth, M. Meijer, M. B. Schilhabel and E. M. H. Wellington. FEMS Microbiology Ecology 13, 13-21, 1993
Molecular Ecology of Filamentous Actinomycetes in soil. Wellington, E.M.H.and P. Marsh. In: Molecular Ecology of Rhizosphere Microorganisms ed. F. O'Gara, D.Dowling and B.Boesten. VCH, Weinheim. 1994. (in press).
Phage-host interactions in soil. Marsh, P.M. and E.M.H. Wellington, FEMS Microbiology Ecology. 1994. (in press).
Actinomycetes Wellington, E.M.H. and Toth, I. K. In: 'Methods of soil analysis', part 2 Chemical and Microbiological Properties -Agronomy Monograph no. 9 (3rd edition) ASA-SSSA. 1994. (in press).

Safety assessment of the deliberate release of two model transgenic crop plants, oilseed rape and sugar beet (BIOT CT-910298)

COORDINATOR:
P. RÜDELSHEIM, Plant Genetic Systems, Gent, B
PARTICIPANTS:
M. RENARD, INRA, Le Rheu, F
W. STIEKEMA, CPRO, Wageningen, NL
P. DALE, IPSR, Cambridge, GB
A. BUCHTER-LARSEN, Danisco, Copenhagen, DK
P. GOUYON , Université Paris Sud, Paris, F
S. WATERS, Monsanto Europe, Brussels, B
J. STREIBIG, RVAU, Frederiksberg, DK
S. JENSEN, NFNA, Horsholm, DK

BACKGROUND INFORMATION

The biotechnological techniques will broaden the range of germplasm available to plant breeders by making accessible, or producing, new gene combinations containing desirable DNA sequences not available from other sources.

The potential and limitations of the new technologies are, as yet, not fully appreciated. Therefore, there is a demand for the precautionary approach reflected in specific regulation. In the EU, Council Directive 90/220/EEC on the deliberate release of genetically modified organisms describes the regulatory framework within which releases can be authorized.

The Organization for Economic Cooperation and Development (OECD) has also provided an international forum for its 24 members (all the major developed countries) to discuss and formulate common approaches and principles.

OBJECTIVES AND PRIMARY APPROACHES

1. Collect and evaluate hard data on pollen dispersal (oilseed rape and sugarbeet).
2. Assess the feasibility of outcrossing and viability of descendants (oilseed rape and sugarbeet).
3. Assess the predictability of transferred gene activity (oilseed rape).
4. Assess the establishment and impact of a new trait introduced into a population (oilseed rape and sugarbeet).
5. Develop and test a computer model for objective 4.
6. Evaluate possible plant instability as a consequence of transformation (oilseed rape).

RESULTS AND DISCUSSION

1. Introduction

This project was based on the results of BAP-project 371 on the 'Study of gene dispersal from plants produced by recombinant DNA'. This project was charac-

terized by a broader scope (including ecological performance and impact), by enhanced designs (moving to larger scale releases) and by focusing on two priority model crops, oilseed rape and sugarbeet.

2. Oilseed rape (OSR)

2.1. Distribution of introduced genes via pollen

Coordinated field trials have been carried out in the U.K., Belgium and France. Typically, a design contained a 'source', with genetically modified plants carrying an easily traceable marker, and a catcher area. Both continuous (i.e. source and catcher area are adjacent to each other) and discontinuous designs have been utilized. In order to reveal rare events, male sterile plants were introduced in some trials. Although the results were dependent on the design, it could be concluded that effective pollen dispersal is limited to the immediate vicinity of each plant.

2.2. Effects of outcrossing to relatives

2.2.1. Performance of plants and populations

Several experimental approaches aimed at estimating the potential as well as the likelihood of gene transfer from oilseed rape to the most prominent European wild (weedy) *Brassicaceae*, including *B. campestris, B. adpressa* (syn. Hirschfeldia incana), *Sinapis arvensis* and *Raphanus raphanistrum*. Experimental steps involved the study of pollination ability, the evaluation of production of F_1 hybrids under field and forced conditions, analysis of the competitive performance of interspecific hybrids and potential for further establishment through reproduction. The results confirm the potential for gene transfer in *B. campestris. B. adpressa, R. raphanistrum* and *S. arvensis* can only transfer genes to oilseed rape under specialized conditions.

2.2.2. Performance of the introduced gene

Many crosses were performed within a large range of genotypes. Observations on the phenotype were used as an indication of the performance of the introduced gene. Markers, such as phosphinotricin tolerance, proved very effective for this purpose. Some lines were identified in which gene expression altered between generations or in different genetic backgrounds. Whenever a change was observed, it resulted in a complete 'silencing' of the introduced gene and a return to the untransformed status.

2.3. Stability of a GM plant

It is known that the insertion site of a transgene in the plant chromosome influences the level of expression. MAR elements (Matrix Associated Regions) are hypothized to decrease variability of gene expression.

Chimeric gene constructs containing a MAR-element (A;TBS), GUS, the lhca*3 promoter and NPTII have been introduced in tobacco. The chicken MAR element (A) induced a 2-3 fold reduction in the variability of transgene expression.

Whereas such approaches may serve future development of stable, high level expression, other developments build intrinsically on instability. Transposon tagging will be used to identify important traits. Oilseed rape plants containing maize transposable elements Ac-Ds have been created. The first observations in the greenhouse have been concluded in order to prepare a regulatory procedure for field release.

2.4. Computer modelling

As this project aimed at building the framework for assessment, modelling took a central role in bringing together the results, in predicting the importance of individual biological parameters and as feedback in designing critical experiments. The dispersal model was at the basis of the design of the last pollen dispersal studies (France & Belgium). Based on the general model incorporating the experimental parameters, simulations can be made on the impact of specific traits.

Sugarbeet

3.1. Dispersal of introduced genes via pollen

Pollen dispersal experiments were designed to measure pollen movement at distances of up to 100 meters, using sugarbeet plants containing a colour marker gene or transgenic plants containing genes that confer tolerance to glyphosate as well as the β-glucuronidase marker gene.

The results were consistent across years (1991-1993) and locations (Belgium and Denmark), and were not dependent on the nature of the pollen source. Using different experimental designs *Beta vulgaris* pollen dispersal was shown to be prominent mainly over short distances (25-35 metres), although pollen could be detected at low frequencies at distances of 100 metres from the pollen source.

3.2. Performance of plants and populations

Using both endogenous marker genes and genes introduced by genetic modification, interspecific hybrids were produced under natural conditions with the wild relatives *B. maritima, B. atriplicifolia* and *B. macrocarpa*. Hybridization with the latter two was less frequent as a result of genetic barriers.

A series of experiments addressed fitness. Competitive ability was evaluated through competition experiments in mixed populations in greenhouse and field.

Potential changes in cold tolerance were investigated in detail. Finally, sensitivity to factors affecting establishment was identified.

The results of the fitness experiments confirm that the traits under investigation did not convey any competitive advantage to transgene sugarbeet.

4. Conclusions

4.1. Based on the practical experience obtained during the project, methods have been developed for the design and monitoring of greenhouse and field trials using genetically modified plants. Methods have been developed to quantify some aspects of the performance of such plants.

4.2. The agronomic performance in the field of the genetically modified plants tested in these projects does not differ significantly from non-modified plants. As with non-modified plants, performance in the field cannot be predicted completely from the performance in the greenhouse. However, greenhouse tests may be valuable in predicting major changes in growth behaviour.

4.3. Relative fitness or weediness of genetically modified crop plants:

The fitness or weediness of genetically modified plants evaluated in the programme were not significantly different from the corresponding non-modified plants.

4.4. Gene dispersal:

a. *Pattern of pollen dispersal:*
Pollen dispersal from plants allowed to flower openly in the field is largely confined to the immediate vicinity of the test field. Effective pollination at long distance is negligible.

b. *Transfer to related species:*
The possibility for transfer of genes to related species is dependent on the particular crop species:
* Although there are weedy relatives of oilseed rape, the transfer of genes has only been established in *Brassica campestris. B. adpressa, Raphanus raphanistrum* and *S. arvensis* can transfer genes to oilseed rape under specialized circumstances.
* Sugarbeet is able to transfer genes.

MAJOR SCIENTIFIC BREAKTHROUGHS

Quantitative data have been generated on the behaviour of genetically modified crops and their offspring. The overview of the data and associated information creates a scientific framework for regulatory decisions. It enables focusing in safety assessment on critical parameters having an effect on the survival and on the physical and genetic dispersal of plants and/or introduced genes in the environment.

MAJOR COOPERATIVE LINKS

Field (and greenhouse) trial work, which obviously was at the core of the project activities, was carried out in different countries and during 2-3 seasons. Logistical requirements included joint design, exchange of materials and exchange of data and results. Additional interaction was at the level of feedback from further laboratory analysis and modelling.

PUBLICATIONS

Individual publications

Arnould J., Gouyon P.H., Lavigne Cl. and Reboud, X., 1993. OGM: une théorie pour les risques. Biofutur pp. 45-50

Chevre A.M., Eber F., Kerlan M.C., This P., 1991. Cytogénétique et Marqueurs moléculaires pour l'analyse des génomes de crucifères. C.R. Acad. Agric. Fr., 17(8): 59-68.

Dale P.J., McPartlan H.C., Parkinson R., Mackay G.R. and Scheffler J.A. 1992. Gene dispersal from transgenic crops by pollen. In The Biosafety Results of Field tests of Genetically Modified Plants and Microorganisms. R. Caspar and J. Landsmann, editors. Biologische Bundesanstalt fur Land-und Forstwirtschaft, Braunschweig, Germany. 73-78.

Dale P.J., Parkinson R. and Scheffler J.A. 1993. Dispersal of genes by pollen — The Prosamo Project. In Opportunities for molecular biology in crop production. D.J. Beadle, D.H.L. Bishop, L.G. Copping, G.K. Dixon and D.W. Hollomon, editors. British Crop Protection Council, Farnham, Surrey, GB. 133-142.

Dale P.J., Irwin J.A. and Scheffler, J.A. 1993. The experimental and commercial release of transgenic crop plants. Plant Breeding, 111:1-22.

Darmency H. 1992. Risk assessment of the release of herbicide resistant transgenic crops: two plant models. Mesogée, 52: 41-42.

Darmency H. 1994. The impact of hybrids between genetically modified crop plants and their related species: introgression and weediness. Molecular Ecology, 3: 37-40.

Darmency H., Lefol E., Reboud X. 1990. Risques associés aux cultures transgéniques résistantes aux herbicides. Résumés du 20ième congrès du groupe français des pesticides, Aiguebelette-Le-Lac, pp. 170-173.

Darmency H., Lefol E., Chadoeuf R., 1992. Risk assessment of the release of herbicide resistant transgenic crop: two plant models. 9ième Colloque international sur la biologie des mauvaises herbes, Dijon, pp. 513-523.
Darmency H. and Renard M. 1992. Efficiency of safety procedures in the experiments with transgenic oilseed rape. 2nd Int. Symp. on the biosafety results of field tests of genetically modified plants and microorganisms. Goslar, R. Casper and J. Landsmann Eds., D!ring Drvck, Braunschweig, 54-60.
Eber F., Chevre A.M., Baranger A., Vallee P., Tanguy X., Renard M., 1994. Spontaneous hybridization between a male sterile oilseed rape and two weeds. Theor. Appl. Genet. (in press).
Fredshavn, J.R., 1993, Modeling plant competition in monocultures of barley and oilseed rape. Acta Agric. Scand., Sect. B, Soil and Plant Sci. 43:38-44.
Fredshavn, J.R., 1993, Use of Substitution Rates to Describe Competition in Mixed Plant Populations. Acta Agric. Scand., Sect. B, Soil and Plant Sci. 43.
Fredshavn, J.R. and Silberg Poulsen G., 1993. Competitiveness of transgenic plants. Proceedings of the Symposium: 'Gene transfer: Are Wild Species in Danger'. Le Louverain, Switzerland, 9 November 1993.
Fredshavn, J.R., The Dynamics of Competition in Mixed Plant Populations. (submitted)
Fredshavn, J.R., Silberg Poulsen G., Huybrechts I. and Rüdelsheim P., 1993. Competitiveness of Transgenic Oilseed Rape.
Kerlan MC., Baranger A., Chevre AM., Eber F., Vallee P., Renard M., Lefol E., and Darmency H., 1992. Outcrossing of transgenic rapeseed to related weedy species. in Bridge: first sectorial meeting on Biosafety, CEC, Wageningen, pp. 32-33.
Kerlan M.C., Chevre A.M., Eber F., Baranger A., Renard M., 1992. Risk Assessment of outcrossing of transgenic rapeseed to related species: I. Interspecific hybrid production under optimal conditions with emphasis on pollination and fertilization. Euphytica, 62:1454-153.
Kerlan M.C., Chevre A.M., Eber F., 1993. Interspecific hybrids between a transgenic rapeseed (*Brassica napus L.*) and related species: cytogenetical characterization and detection of the transgene. Genome, 36: 1099-1106.
Lefol E., Danielou V., Darmency H., Kerlan M.C., Chevre A.M., Renard M., Reboud X. 1992. Escape of engineered genes from rapeseed to wild Brassiceae. in Complexes d'especes, flux de genes et ressources genetiques des plantes, BRG ed., Lavoisier Tec. et Doc., pp 616-617.
Lefol E., Danielou V., Darmency H. 1994. Predicting hybridization between transgenic oilseed rape and wild mustard. Agriculture, Ecosystems and Environment (in press).
Lefol E., Danielou V., Darmency H., Boucer F., Maillet J. and Renard M. 1994. Gene dispersal from transgenic crops: I Growth of interspecific hybrids between oilseed rape and the wild hoary mustard. (submitted).
Lefol E., Danielou V., Darmency H., Kerlan M.C., Vallee P., Chevre A.M., Renard M., Reboud X. 1991. Escape of engineered genes from rapeseed to wild Brassiceae. Proceedings. Brighton Crop Protection Conference — Weeds-, 3, pp. 1049-1054.
Lefol E., Darmency H. 1993. Risk of interspecific gene transfer from transgenic canola to wild relatives. Abstracts: Eighth Crucifer Genetic Workshop, Saskatoon, SK, Canada, pp. 56.
Lefol E., Fleury A. and Darmency H. Gene dispersal from transgenic crops: II Hybridization between oilseed rape and the hoary mustard. (in preparation)
Lefol E., Fleury A. and Darmency H. Genetic structure of a wild mustard population (in preparation)
McPartlan H.C. and Dale P.J. 1994. An assessment of gene transfer by pollen from field grown transgenic potatoes to non-transgenic potatoes and related species. Transgenic Research. (in press).
Mlynarova L., Loonen A., Keizer P., Stiekema W.J. and Nap J.P., 1994. Reduced variation of transgene expression conferred by the chicken lysozome MAR sequence. Poster, Plant Genome II, San Diego, USA, January 4-8.
Mlynarova L., Loonen A., Heldens J., Jansen R.C., Keizer P., Stiekema W.J. and Nap J.P., 1994.Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. The Plant Cell 6, 417-426.
Mlynarova L., Loonen A., Jansen R.C., Stiekema W.J. and Nap J.P., 1994. A matrix-associated region reduces position effects in transgenic plants. Abstract, ISPMB, Amsterdam, Netherlands, June 20-25.

Mlynarova L., Loonen A., Jansen R., Keizer P., Stiekema W.J. and Nap J.P., 1993. Reduced variation of transgene expression conferred by the chicken lysozyme MAR sequence. Poster EEC TEMPUS workshop 'Gene manipulation for Crop Improvement', Ceske Budejovice, Czech Republic, October 25-29.
Nap J.P., Mlynarova L., Keizer P., Loonen A., and Stiekema W.J., 1992. The influence of Scaffold Attachment Regions (SARs) on the predictability of the expression of transgenic DNA in plants. Abstract (p. 34) and Poster BRIDGE 1st Sectoral Meeting on Biosafety, Wageningen, December 6-9.
Nap J.P., (1993). How to analyze data sets with non-parametric statistical tests. CPRO-DLO, 52 pp.
Nap J.P., Mlynarova L. and Stiekema W.J., 1993. From transgene expression to public acceptance of transgenic plants: a matter of predictability. Discussion paper OECD Workshop 'Multidisciplinary Assessment of Transgenic Crop Plants: Technical, Agronomic and Ecological Aspects', Chania, Crete, Greece, 8 pp.
Nap J.P., Keizer P. and Jansen R., 1993. First-generation transgenic plants and statistics. Plant Mol. Biol. Rep. 11, 156-164.
Nap J.P., van Spanje M., Dirkse W.G., Baarda G., Mlynarova L., Loonen A., Grondhuis P. and Stiekema W.J., 1993. Activity of the promoter of the Lhca3.St.1 gene, encoding the potato apoprotein 2 of the light harvesting complex of Photosystem I, in transgenic potato and tobacco plants. Plant Mol. Biol. 23, 605-612.
Nap J.P., Mlynarova L., Loonen A. and Stiekema W.J., 1994. Position-independent expression of the GUS reporter gene in the presence of a MAR element. Abstract, Plant Genome II, San Diego, USA, January 4-8.
Scheffler, J.A., Parkinson R. and Dale P.J. 1993. Frequency and distance of pollen dispersal from transgenic oilseed rape (*Brassica napus*). Trans Res. 2:356-364
McPartlan H.C. and Dale P.J. 1994. The transfer of introduced genes from field grown transgenic potatoes to non-transgenic potatoes and related solanaceous species. Trans Res. (in press).
Scheffler J.A., Parkinson R. and Dale P.J. 1994. Evaluating the effectiveness of 200 and 400 metre isolation distances for oilseed rape (*Brassica napus*) field plots. J. Exp. Bot. (in preparation).
Scheffler J.A. and Dale P.J. 1994. Opportunities for gene transfer from transgenic oilseed rape (*Brassica napus*) to related species. Transgenic Research. (in press).

Stability, genetic transfer and ecology of fungi used as biocontrol agents (BIOT CT-910290)

COORDINATOR:
G. RIBA, INRA, Guyancourt, F
PARTICIPANTS:
Y. BRYGOO, Univ. Paris XI, Orsay, F
D. RHODES, ZENECA Agrochemicals, Bracknell, GB
L. GULLINO, DI.VA.PRA., Univ. Torino, I
J. COREMANS-PELSENEER, Univ., Bruxelles, B

BACKGROUND INFORMATION

Uses of filamentous fungi as biological control agents are increasing and are highly diverse with respect to the targets (insect pests, weeds, fungal plant pathogens), the strategies (inoculative treatments or inundative treatments), the partners (public research institutions, private companies, farmers consortiums, plant protection service, etc.). The efficacy of fungal preparations has been proved many times by consistent field experiments since commercial products have been introduced in USA, Europe and several developing countries. For example, in 1993, two fungal biopesticides based on *Beauveria* strains have been registered in France, and several other preparations are being evaluated in Italy, France, Belgium and the Netherlands.

To increase the agronomic use of filamentous fungi in the crops it is necessary:
— to enhance their efficacy by protective formulations and strain improvement through parasexual cycle or transformation;
— to assess the risks, following the framework of the European Guidelines 90/220 and 414, i.e safety to humans and mammals; safety to other organisms; safety in the environment;
— to evaluate the possibilities of genetic exchanges with the resident microflora.

OBJECTIVES AND PRIMARY APPROACHES

Researches done in this project were focused on:
(i) main sources of genetic plasticity in fungi,
(ii) potential risks induced by the use of recombinants or transformants,
(iii) methodologies to monitor strains in the environment.

Two genera representative of the diverse uses of fungal biocontrol agents have been selected: non pathogenic strains of *Fusarium oxysporum* can be applied in the rhizosphere in order to compete with indigenous plant pathogenic strains of the same species, a highly virulent strain of *Beauveria bassiana* 147 has been registered as a BCA against European corn borer, while another one, *Beauveria brongniartii* 96 has been introduced and acclimated in La Réunion island against *Hoplochelus marginalis*, a white grub of sugar cane.

At the beginning of the project very few techniques had been validated for characterization of the DNA of such species. For this reason, an important contribution of the project has been to provide molecular tools to study transformation and subsequent genetic stability of strains, and to track them under natural conditions.

RESULTS AND DISCUSSION

I — Genomic typing

The genetic peculiarities of the selected fungal genera are
(i) non sexual cycle,
(ii) a genome size which is 4 times larger than a yeast genome and about 10 times larger than that of bacteria.

Strains of *Fusarium* and *Beauveria* have been characterized by RFLP, RAPD, CHEF, sequencing and hybridization applied to the whole or part (rDNA, a gene region) of the genome.
The relationships between pathogenic and non pathogenic *Fusarium* strains have been analysed using a sample of 60 pathogenic isolates belonging to 12 different formae speciales together with 36 non pathogenic isolates. The resulting polyphyletic tree showed that each phylum groups together isolates from different formae speciales while, isolates belonging to the same f.sp. are contained within two phyla indicating that the specialization to the same host plant can result from independent events. Genetic separation of non pathogenic strains is correlated with their ecological origin.

In the same way, comparison of restriction patterns from amplified ribosomal internal transcribed spacers (ITSs) allowed separation of 32 *Beauveria brongniartii* strains into 6 haplotype groups. One of them contains all strains isolated from the white grub *Hoplochelus marginalis*. A similar genetical grouping of strains isolated from the same host insect was observed by comparison of RAPD patterns of strains of *B. bassiana* pathogenic towards the corn borer *Ostrinia nubilalis*. Moreover, PCR-RFLP analysis of the 28S ribosomal sub-unit in *B. brongniartii* revealed the presence of insertional elements of about 350-450 bp. Cloning and sequencing of one of these insertions revealed it to be typical of group I introns. By studying the distribution of these group I introns in various strains isolated from *H. marginalis* it was possible to discriminate one of these strains, the biological agent Bt 96, and so to specifically monitor its presence in the environment.

Molecular typing of this sort is particularly useful since probabilities of genetic exchanges are related to the genetic distances between strains.

II — Sources of genetic instability

Hyphomycetes may possess particular sources of genetic instability since exchange of genetic information may occur during cell fusion. Somatic recombination through parasexual cycle has already been described in several fungal genera. In addition, exchanges of cytoplasmic factors may occur during the parasexual cycle, while transposition may be associated with the presence of transposable elements.

A) Parasexual cycle

The parasexual cycle of *Beauveria bassiana* and *Fusarium oxysporum* has been described previously. Following hyphal or conidial anastomosis cytoplasm is combined and a heterokaryotic stage occurs; eventually diploids are formed by nuclear fusion and they finally will yield parental or recombinant types after haploidisation during hyphal growth or conidiogenesis. The low frequency of fusion events may be increased by protoplast fusion.

Interspecific protoplast fusions have been obtained between two closely related species *B. bassiana* and *B. sulfurescens*. It has been proved that some of the rare products of fusion are stable during mitotic cycles though they have a partially diploid structure. Secondly some of those recombinant strains acquired some

characteristics of the parental strains and this may confer a selective advantage during few steps of the life cycle.

Two types of natural barriers for fusion exist either at the cell wall level or within the cytoplasm. Interstrain fusions may not be possible if the parental strains do not belong to the same vegetative compatible group (VCG). On an other hand, we have demonstrated that strains of *B. bassiana* belonging to the same VCG can exchange genetic informations when they are introduced by injection into the host insect.

B) Transposable elements

1) *Identification of transposable elements:* the high level of genetic variation in *F. oxysporum* led us to postulate that a part of the variability observed in this species might result from the activity of transposable elements. Representatives of both DNA transposons and retrotransposons have been identified. Firstly, the characterizaton of repetitive dispersed sequences has allowed the identification of elements with characteristics of LTR-retransposons (*Foret* and *palm*). These elements are inactive but offer a powerful tool to characterize strains. Secondly, the molecular analysis of spontaneous mutants in the *nia* gene encoding nitrate reductase led to the identification of at least four families of transposable elements (*Fot1, Fot2, FML* and *Hop*). All these families resemble bacterial transposons and contain active copies.

2) Mutagenic effects of transposable elements: it has been demonstrated that the activity of bacterial-like elements determines a high level of genetic plasticity i.e., inactivation of the target gene through insertion, diversification of the nucleotide sequence through imprecise excision and probably chromosomal rearrangements as suggested by the extensive karyotype variation observed among field isolates. All these elements are widely dispersed in the species *F. oxysporum* and their mutagenic effect may provide a mechanism for generating variability that may be important in the appearence of virulence factors.

3) *Potential for gene transfer:* in order to test the occurrence of a possible gene transfer, the presence of active transposons was examined because these elements have built-in mechanisms for excision and reintegration and thus appeared the elements best equipped to invade a genome. First, the distribution of *Fot1* elements was analysed in different species of the *Fusarium* group whose phylogenetic relationships are well established. The discovery of *Fot1* copies in the distant species *F. solani*, which are highly homologous to the *Fot1* element isolated from *F. oxysporum*, suggests that a horizontal transfer between these species may have occurred recently in nature. The comparison of their nucleotidic sequences would confirm such an event.

At the initiation of the project, the only transposon reported in fungi was the retroelement Tad from *Neurospora crassa*. In the course of the project, at least six families of transposable elements have been identified in the species *Fusarium oxysporum*. Some of them share a bacterial-like structure and therefore correspond to a new class of fungal transposable elements. In addition, the strategy developed in *Fusarium*, that is the screening of unstable mutants in the gene encoding nitrate reductase, has been applied to *Beauveria* and has allowed the identification of different inserted sequences with features of active transposable elements. This approach appears the most effective way to screen active transposons and the use of the *nia* gene as a target seems to be of general application. Other aspects of the studies carried out in these fungi such as the mutagenic effect of the trans-

posable elements and the possibility that genetic information can transfer horizontally, should provide a model for studies in other species.

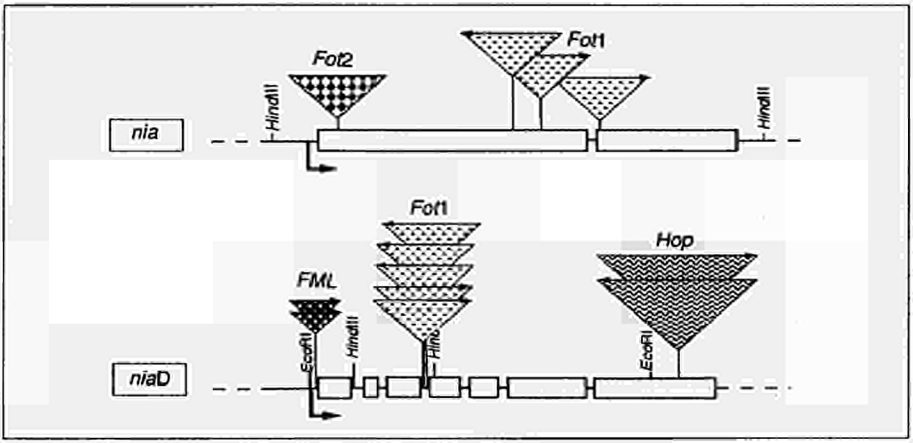


Fig. 1: Identification of four families of bacterial like transposable elements in Fusarium oxysporum *using genes encoding nitrate reductase,* nia *D gene from* Aspergillus nidulans *and* nia *gene from* F. oxysporum *as insertion traps.*

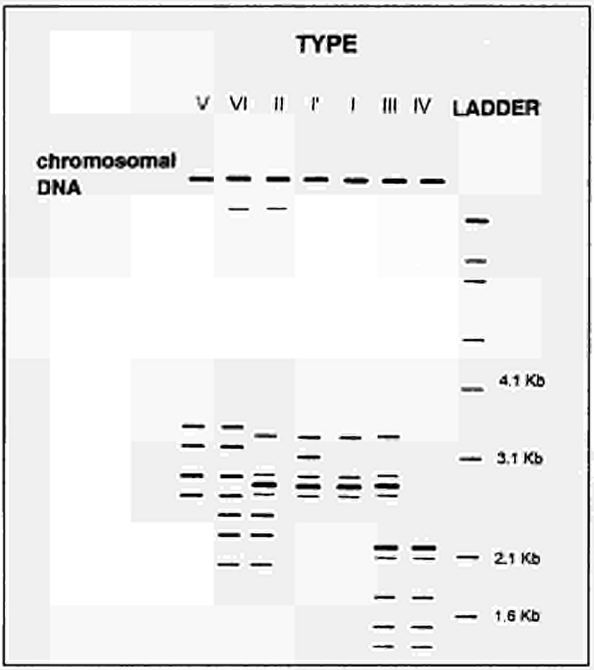


Fig. 2: Presence of dsRNA elements in isolates of Beauveria bassiana*: diagram showing the patterns of bands of each of the dsRNA types; the relationships between the patterns can be seen, and the size of the bands estimated by comparison with the ladder.*

3) ds RNA

A collection of 84 *B. bassiana* strains, isolated from widely-dispersed field sites was examined for the presence of self-replicating extrachromosomal elements. No

evidence was obtained for the presence of mitochondrial plasmids, but 78% of strains were found to contain double-stranded RNA (dsRNA) elements. This is the first description of such elements in the species *Beauveria*; the elements were classified into seven main groups, although some of these groups may represent combinations of two or more elements in the same strain. One pattern was found to be widespread geographically, and represented 70% of those strains which contained dsRNA elements. Since the literature suggests that fungal dsRNA elements are non-infectious except by hyphal anastomosis, such elements should prove to be a valuable tool for determining the frequency of hyphal fusion and exchange of extrachromosomal elements between strains of *Beauveria*.

III — Transformation of fungal strains

As already described, protoplast fusion is a potentially powerful technique for strain improvement since different genes may be included in one strain. Nevertheless in order to introduce a specific characteristic, transformation is an alternative process with a great potential.

Transformation of fungi requires plasmid integration. Different vectors have been constructed with strong promoters, and several marker genes have been need. Cotransformation is a more efficient system where two plasmids are simultanously introduced one with a marker gene the other with the gene of interest. Frequency of transformation is typically high (approximately 500 transformants/µg DNA). Several transformants of *Fusarium oxysporum* and *Beauveria bassiana* have been obtained with heterologous DNA carrying *gus A* , *ben A, nia D* and *ATP* sulphurylase genes. ATP sulfurylase gene (SC) and nitrate reductase gene (nia D) from *Aspergillus nidulans* were introduced into *Beauveria bassiana* in order to assess their stability in the presence and absence of selection pressure. When counter selection against the foreign DNA was applied, occasionnal revertants appeared; they were typically associated with loss of the transforming DNA. In the absence of selection pressure, transformants remained stable, even after three months of serial subculture.

Transformation with homologous DNA requires the cloning of genes; the *nia gene* of *Beauveria* was cloned by probing with products of PCR amplification using degenerate nucleotide the sequences of which were deduced from known *nia* gene; the ATP sulphurylase gene of *Beauveria* has been cloned by applying the strategy previously described.

IV — Genome plasticity under natural environments

In non sterile soils *F. oxysporum* and *B. bassiana* strains have to compete with natural microflora and microfauna. They are able to survive in those conditions since soil is a 'natural reservoir' but the introduced population is not always sufficient to suppress the pest population. Among exceptions, some strains of *B. brongniartii*, such as *B. brongniartii* 96, are able to control the pest population for several years in sugar cane fields of La Réunion island. Characterization of the behaviour of entomopathogenic fungi in soil is complicated by the fact that total colony counts on selective media fail to distinguish between the infective propagules (conidia) and other forms of fungal biomass. An original technique based on heat treatment was therefore designed which eliminates hyphal biomass while retaining the viability of conidia in the soil. This allow conidiation and the persistence of conidia in the soil, to be monitored directly.

In sterile soils *Fusarium* and *Beauveria* have the potential to reach a plateau from 10^7 to 10^8 propagules/g depending on the physical and chemical characteristics of the soil.

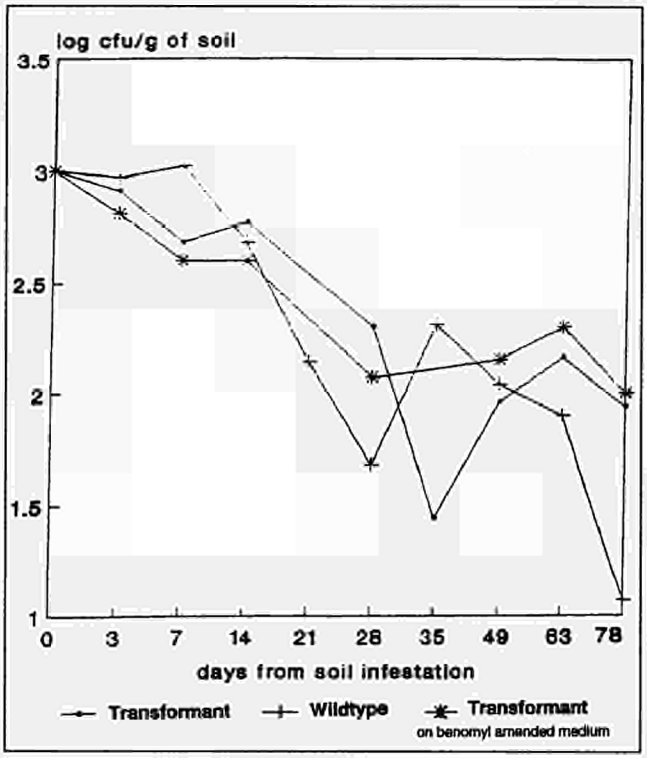


Fig. 3: Survival of benomyl resistant transformant and benomyl sensitive F. oxysporum *in soil microcosms.*

The distribution of selected strains of antagonistic Fusarium spp. was tracked using molecular tools: both the electrophoretic karyotype patterns and the random amplified DNAs were stable after six months from the introduction in natural and sterile soils and allowed the unequivocal recognition of the introduced strains.

Genetic transformation to benomyl resistance does not affect the survival of *Fusarium* strains, unless benomyl is added to the soil. No significant effects on natural microflora present in the soil and rhizosphere were observed after the addition of either transformed or non manipulated antagonists.
A benomyl resistant transformant and a benomyl-sensitive *F. oxysporum* marked by orange pigmentation were introduced into soil microcosms in the presence or absence of test plants, in order to evaluate the possibility of genetic transfer of benomyl resistance under natural conditions. Experiments carried out for three or six months demonstrated that orange colonies were never isolated on benomyl-amended PDA. Population densities slowly decreased during the experiment, and were not influenced by the presence of test plants.

On leaf surfaces, *Beauveria* is not able to survive beyond 4 days after the application without formulation. It is highly susceptible to UV, which does not appear to induce any genetic modification.

MAJOR SCIENTIFIC BREAKTHROUGHTS

At the onset of the project, no background was available on the genetic ecology of filamentous fungi. At the conclusion of this project, we are able to underline three main points:

— Techniques and methods which could be used by applicants to release and register fungal biocontrol agents have been developed: molecular typing techniques, highly informative probes, and transformation systems with several reporter genes.

— Secondly, original scientific data have been generated on the stability of introduced DNA, and the presence of transposons and dsRNA elements in biocontrol fungi. These merit further investigations. Further work should concentrate on main problems:
  (i) are the active transposable elements able to horizontal transfer?
  (ii) have the active transposable elements a practical application in biotechnology?
  (iii) how to monitor polyclonal artificial populations?

— Finally, results of the project allow us to put forward recommendations for the assessment of risks induced by recombinant fungal procedures for agronomic purposes:
  (i) to estimate the probability of genetic exchanges between the introduced strains and those already present in the target environment (VCG, genetic distances, protoplast fusion),
  (ii) to research active transposable elements,
  (iii) to monitor introduced fungal populations.

MAJOR COOPERATIVE LINKS

Joint meetings and research visits have proved especially valuable for exchange of ideas, comparison of data and evaluating progress. Exchanges of improved strains, plasmids, cloned genes, methods and protocols were performed between the five laboratories participating in the project, and a training period (april 93) was performed by Q. Migheli (Univ. Torino) at the University of Orsay, F.

PUBLICATIONS

Individual publications

Couteaudier, Y., Maurer, P., Viaud, M., Riba, G. — 1994. Genetic stability of fungi. VIth Intern. Coll. on Invertebr. Pathol. and Microbial Control. Montpellier, France. August 28 — 1 Sept. 94.

Langin, T., Capy, P., Daboussi, M.J. — The transposable element, Impala, a fungal member of the Tcs-mariner superfamily. Mol. Gen. Genet. (in press).

Mezzalama M., Migheli Q., Gullino M.L., Garibaldi A. — 1993. Dinamica di popolazione di *Fusarium oxysporum* antagonisti geneticamente manipolati in microcosmi tellurici. Micologia Italiana, 3: 30-35.

Mezzalama M., Mocioni M., Gullino M.L. — 1994. Survival of antagonistic *Fusarium* spp. in soil. Microbial Releases, 2 (in press).

Migheli Q., Berio T., Gullino M.L. — 1993. Electrophoretic karyotypes of *Fusarium* spp. Experimental Mycology, 17: 329-337.

Migheli Q., Berio T., Gullino M.L., Garibaldi A. — 1994. Electrophoretic karyotype variability among pathotypes in *Fusarium oxysporum* f.sp. dianthi. Plant Pathology.

Migheli Q., Cavallarin L. — 1994. Characterization of antagonistic and pathogenic *Fusarium oxysporum* isolates by random amplified polymorphic DNAs (RAPDs). Molecular Biotechnology (in press).

Migheli Q., Mezzalama M., Friare O., Gullino M.L. — 1994. Risk analysis for wild type and genetically engineered antagonistic *Trichoderma* and *Fusarium* spp. IOBC Bulletin (in press).

Migheli, Q., Mezzalama, M. Friard, O., Gullino, M.L. — 1994. Risk analysis for wild type and genetically engineered antagonistic *Trichoderma* and *Fusarium* spp. IOBC Bulletin (in press).

Neuveglise, C., Brygoo, Y. — Identification of group 1 introns in the 28S rDNA of the entomopathogenic fungus *Beauveria brongniartii*. Current Genetics. (in press).

Neuveglise, C., Brygoo, Y., Vercambre, B., Riba, G. — 1994. Comparative analysis of molecular and b iological characteristics of strains of *Beauveria brongniartii* isolated from insects. Mycol. Res., 98 (3), 322-328.
Osborne, D.J., Rhodes, D.J., Smith, J.D. — 1994. Presence of double stranded RNA elements in isolates of *Beauveria bassiana* and their association with virulence. IId European Conference of Fungal Genetic. Luteren. The Netherlands, April 28- 2 May 94.
Rakotonirainy, M.S., Cariou, M.L., Brygoo, Y., Riba, G. — 1994. Phylogenetic relationships within the genus *Metarhizium* based on 28S rRNA sequences and isozyme comparison. Mycol. Res., 98(2), 225-230.
Tillemans, F., Lannoye, R., Coremans-Pelseneer, J. — 1993. Techniques for observation of fungi in soil and fungal effects on plants. Med. Fac. Landbouww. Univ. Gent, 58/ b: 1229-1242.

Genetic tools for constructing GEMs with high predictability in performance and behavior in ecological microcosms, soils, rizhospheres and river sediments (SSMA) (BIOT CT-910293)

COORDINATOR:
J.L. RAMOS, C.S.I.C.-E.E.Z., Granada, E
PARTICIPANTS:
V. de LORENZO, C.S.I.C.-C.I.B., Madrid, E
S. MOLIN, T.U.D., Lyngby, DK
K.N. TIMMIS, GBF-I, Braunschweig, D
F. O'GARA/D. DOWLING, U.C. Cork, IRL.
I. WAGNER-DÖBLER, GBF-II, Braunschweig, D

BACKGROUND INFORMATION

A large proportion of organic compounds of biological and chemical origin are mineralized, i.e. converted to carbon dioxide, water and other inorganic compounds, predominantly by microbes, as part of the continuous cycling of carbon between inorganic and organic states. However, many xenobiotics (man-made chemicals) are recalcitrant and persist in the environment because they have novel structures or substituents not found in biogenic molecules, are chemically very stable, are toxic for microorganisms, or inhibit degradative pathways. In these cases the evolution of new activities in the laboratory may be helpful because the frequency and types of genetic events needed (mutations, alteration of gene expression, gene dosage and gene transfer, etc) can be optimally achieved under powerful selective conditions. Several strategies have been applied by the research groups participating in this EC BRIDGE program to construct bacteria able to eliminate a number of chlorinated aromatics such as polychlorinated biphenyls (PCBs), and to improve the biodegradative properties of bacteria able to eliminate organic solvents. These developments have been coupled to conditional gene expression systems involving two aspects: recombinant organism containment, which includes the assembly of lethal genes in regulatory networks for the mineralization of aromatic hydrocarbons, and recombinant gene containment, through the establishment of barriers to lateral gene transfer.

These developments allowed the construction of GEMs that display a high degree of functional and ecological predictability, which has been validated in microcosms of agricultural soils and river sediments.

OBJECTIVES AND PRIMARY APPROACHES

1. Construction of PCB degraders

PCBs, a group of compounds that have been introduced into the environment as a result of industrial activity, have proven recalcitrant to microbial degradation under natural conditions. They can enter the food chain and eventually accumulate in lipid-rich tissues in humans and animals. Because of the health hazards associated with PCBs, removing them from the environment is a high priority task. Since PCBs are often found in large volumes of soils and sediments, albeit at low concentrations, in situ degradation may be an effective bioremediation approach. Inoculants are therefore needed which can survive and degrade-PCBs under environmental conditions. Although a number of selected strains are available

which aerobically degrade a wide spectrum of analogs, none of them has been successful as a bioremediation agent. The problems identified with aerobic in situ degradation of PCBs occur both at a biochemical level (co-metabolism, insufficient regulation, incomplete mineralization) and an ecological level (insufficient survival and low activity of the target strains in contaminated sites). Below we describe our strategy to overcome these obstacles.

The catabolic genes *bph*ABCD of *Pseudomonas* sp. LB400, the best known PCB degrader, were cloned and the regulatory circuits dissected. The operon was cloned into a mini-Tn5 vector developed in BRIDGE program. The *bph* cassette was transferred to the chromosome of indigenous microorganisms able to establish in target ecological niches, namely, the sugar-beet rhizosphere and sediments from the River Spittelwasser, a tributary of the River Elbe. A screening strategy was then set up to select microbes able to degrade the chlorinated benzoates resulting from metabolism of PCBs.

2. Strain containment and rDNA containment.

Since the ecological behavior of microorganisms introduced into the environment has not so far been extensively documented, it seems to be useful to restrict their survival in order to increase their predictability in the environment. Constructions with a high degree of functional and ecological predictability, and equipped with barriers to lateral gene transfer, can be designed in the laboratory by using lethal genes coupled to the regulatory system of the pollutant to be removed. As a model system we have used *Pseudomonas* bacteria able to use aromatic organic solvents as the sole carbon and energy source.

Construction of the contained GEMs was based on the *gef* killing gene coupled to regulatory networks for the degradation of toluates. The system is based on two elements: the control element, which consists of a fusion between Pm, the promoter of the pathway for metabolism of toluates, and a gene encoding a repressor protein (in our case the LacI protein), and the xylS gene encoding the positive regulator of the Pm promoter. The other element, the killing cassette, consists of a fusion between the Plac promoter and the *gef* gene, which encodes a killing function. In the presence of an aromatic compound such as toluate, the bacteria produces LacI protein, which prevents expression of the killing function. In the absence of these effector compounds, expression of the killing cassette is no longer prevented, and the bacteria dies (Figure 1).

To limit the rate of lateral transfer of rDNA, a gene containment system was conceived. This system also involves two elements: the killing element, which consists of the *col*E3 gene encoding for colicin E3, an RNase that cleaves the 16S rRNA of all prokaryotes analyzed so far and the control element, which consists of the immE3 gene which codes the ImmE3 protein, a specific repressor of the lethal function. In the GEM the gene specifying the lethal function is closely linked to the rDNA determining the novel trait, whereas the immunity function is located unliked to the new trait. Transfer of the determinant of the novel phenotype to a new organism will be accompanied by transfer of the lethal gene, but not by that of the immunity gene and, as a consequence, the recipient organism will be killed (Figure 2).

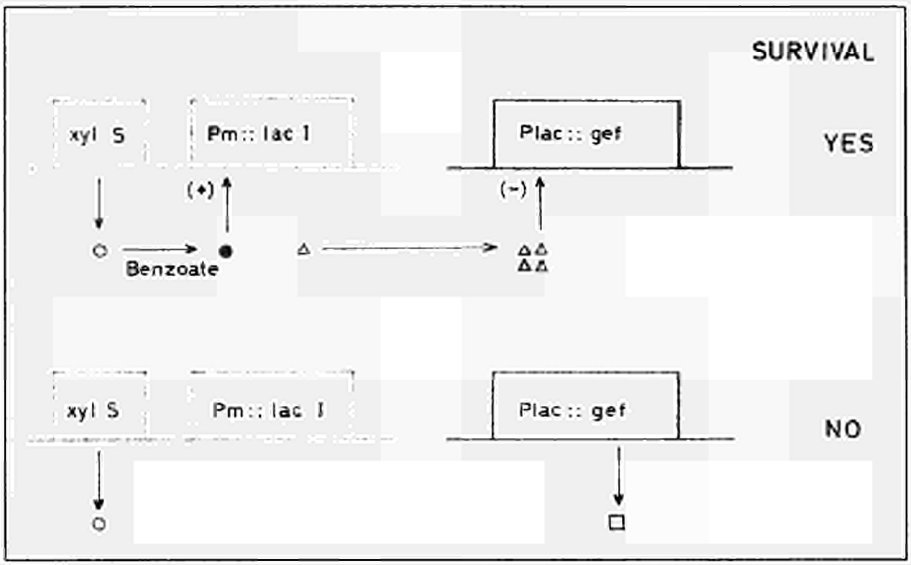


Fig. 1: Programmed survival and death of Pseudomonas *putida based on the presence or absence of a pollutant. The biological containment system is an example of these designed to control survival of bacteria by modifying the availability of nutrients in this case, benzoate.*

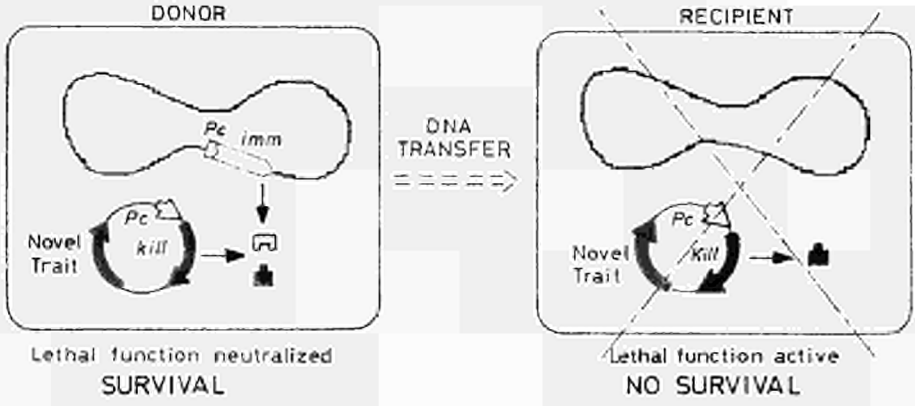


Fig. 2: Gene containment system based on a lethal donation of a killing function. Barriers to gene transfer were designed to avoid the establishment of rDNA in hosts other than those specifically designed as recipients.

RESULTS AND DISCUSSION

1. Degradation of PCBs: Monitoring responsiveness of PCB-degrading operons to pathway substrates and transfer of PCB-degrading genes to target microbes. *Pseudomonas* sp. LB400 uses biphenyl as the sole C and energy source and is able to co-metabolize a wide range of polychlorinated biphenyls to chlorobenzoates.

Expression of the *bph* genes is driven by promoters located at the 5' end of the operon. Because responsiveness of the regulatory circuit to pathway substrates can be one of the major bottlenecks in the efficient performance of PCB-degraders, we have examined in detail the regulation of the system *in vivo*, with particular regard to its potential to express *bph*-encoded enzymes under different conditions, and its inducibility by biphenyl. To reproduce the system faithfully *in vivo*, all the elements involved in the regulation of the system were incorporated into hybrid mini-Tn5 transposons carrying fusions of *'lacZ* to the two first encoding sequences (*orf*0 and *bph*A). We then analyzed the activity of the *bph*::*lac*Z fusions in response to pathway substrates and other aromatics in the homologous genetic background of *Pseudomonas* sp. LB400, and in the heterologous background provided by *Pseudomonas putida* 2442. Our results showed that regardless of the presence of biphenyl, the *bph* genes are expressed constitutively at relatively high levels, and that exposure of the cells to biphenyl results in a moderate increase over these already high basal levels. The induction of the system required the integrity of the orf0 and its pattern was found to be growth phase-independent and relatively insensitive to the presence of rich carbon sources. These results suggest that operation of the bph genes, although not well regulated, can occur under different environmental conditions.

On the basis of the above results, a mini-Tn5 bearing the whole *bph* operon was constructed and transferred to *Pseudomonas* F113, a biocontrol strain isolated from the rhizosphere of sugar-beet, and to seventeen gram-negative bacteria isolated from the sediment of the River Spittelwasser. Integrety of the inserted operon in the recipient bacteria was monitored by hybridization using the *bph*C gene as a probe, or by PCR. Functionality of the operon in the target microbes was confirmed by their ability to grow on biphenyl. The novel genetic trait was stably maintained in the recombinant microbes, as determined by the fact that the whole cassette was kept intact after growth of the strain for more than 200 generations in the absence of selective pressure. The *bph* operon integrated on the host chromosome was not mobilized to related microbes.

The recombinant microbes retained wild type characteristics regarding traits important for their ecological survival, i.e., production of siderophores and of the fungicide phloroglucinol. *Pseudomonas* F113bph was able to colonize plant roots in nonsterile soil, and was as competitive as the wild type in co-inoculation tests.

2. Development of biologically contained strains and barriers to lateral transfer of rDNA

2.1. Analysis of the control elements of the TOL catabolic pathways and design of engineered regulatory cascades

The TOL plasmid pWW0 encodes a pathway for the mineralization of toluene/xylenes via benzoate/toluates. The genes are organized in two operons, the 'upper' pathway operon that encodes the proteins for transformation of aromatic hydrocarbons into benzoates, and the *meta* operon for metabolism of the latter to Krebs' cycle intermediates. The XylR protein with toluene stimulates transcription from the 'upper' pathway operon promoter and from the *xyl*S gene, whose gene product in turn stimulates transcription from the meta operon promoter (Pm). The XylS regulator also stimulates transcription from Pm in response to benzoate effectors. Because the TOL system was chosen for the coupling of the regulatory network to the killing function in order to develop a biological containment system (see below), we were interested in elucidating the circuits controlling the expression of this regulatory system. Accordingly we analyzed time response to effectors, steady-

state levels of mRNA expression under different growth conditions, and catabolite control in this system. The ultimate regulator of the pathways, the XylR protein was found to be made constitutively and its production to be self-regulated. We found that bacteria growing on minimal medium responded immediately to the addition of effectors. The exposure of cells to an 'upper' pathway substrate such as *m*-xylene resulted in full expression from the 'upper' pathway operon promoter (Pu) and the *xyl*S gene promoter (Ps) in less than 2 min, with expression from the Pm promoter requiring about 5 min, which accounted for the time required to attain maximun level of XylS protein in the cell. Bacteria exposed to benzoate effectors exhibited high levels of only *meta* cleavage pathway mRNA. High mRNA levels were maintained with time until pathway substrates were exhausted. In bacteria growing in rich medium, the response of induction of the *meta* cleavage pathway to benzoate effectors was also immediate, whereas in this medium the response of the Pu and Ps promoters to 'upper' pathway effectors was delayed by hours. We found that amino acids present in the rich medium exerted catabolite repression of these two promoters.

The potential of the regulatory elements of the TOL system has been exploited to control gene expression through artificial cascades. A general method was developed, based on the ability of the T7 RNA polymerase (T7pol) to drive transcription from a specific promoter (P_{t7pol}). A specialized mini-Tn5 *xyl*S/Pm::T7pol was constructed, which contains the structural gene for T7 polymerase downstream of the XylS/Pm promoter of the TOL *meta* cleavage pathway. This transposon was stably transferred to the chromosome of *P. putida*. A series of P_{t7pol}::reporter genes (*lux, lac,* etc) were also inserted in the host chromosome via an independent mini-Tn5 element. In this way, genes downstream of the T7 promoter are expressed in response to the exposure of the cell to aromatic compounds. This type of regulatory cascade is now being exploited to control expression of catabolic pathways that cannot be precisely regulated by natural systems.

2.2. *Behavior of contained* P. putida *strains.*

A model substrate-dependent suicide system to biologically contain recombinant *Pseudomonas putida* was conceived. The system consists of two elements: One element was made from a fusion between a synthetic lac promoter ($P_{A1-04/03}$) and the *E. coli gef* gene, which encodes a membrane protein that collapses the membrane potential and leads to cell death. This element is contained in a mini-Tn5 transposon so that it can be integrated randomly on the *Pseudomonas* (or any other gram-negative) bacterial chromosome.

The second element is harbored in the wide-host range plasmid pCC102, and is designed to control the killing element. It bears the promoter of the TOL plasmid-encoded *meta*-cleavage pathway operon (Pm) and the *lac*I gene, encoding the *lac*I repressor, plus *xyl*S2, encoding for a positive regulator of Pm. In liquid cultures under optimal growth conditions and in sterile and nonsterile soil microcosms, *P. putida* (pWW0) bearing the containment system behaved as predicted. In the presence of a XylS effector such as m-toluate, the LacI protein was synthesized, preventing expression of the killing function. In the absence of effectors, expression of the $P_{A1-04/03}$::*gef* cassette was not longer prevented and a high rate of cell killing was observed. Fluctuation tests showed that mutants resistant to cell killing arose at a frequency of around 10^{-5} to 10^{-6} per cell per generation. Mutations were linked to the killing element rather than the regulatory one. In bacteria bearing two copies of the killing cassette, the rate of mutants resistant to killing decreased to as low as 10^{-8} per cell per generation.

A streptomycin-resitant derivative of the above strain, called *Pseudomonas* sp. 4.11, was isolated as more sensitive to killing by the *gef* gene, and the rate of mutants was found to be at least one order of magnitude lower. The killing cassette in P. putida sp. 4.11 was inserted on the TOL plasmid rather than on the chromosome. When this bacterium was introduced in soil microcosms, it was not able to establish in soils unless they had been supplemented with m-toluate (0.01%, wt/wt). The introduction of the contained bacteria did not have significant effects on natural populations present in the soil and able to use p-hydroxyphenylacetic acid or benzoic acid as carbon sources.

2.3. A new recombinant gene containment system.

The transfer and spread of new traits from a GEM to indigenous microbiota is, in principle, an undesired event, because the ecological consequences are unknown. For this reason we designed barriers to reduce the rate of horizontal gene transfer. Colicin E3 is a bacteriocin produced by certain colicinogenic strains of *Escherichia coli*, which kills microorganisms by inhibiting protein synthesis. It does this by specifically cleaving a fragment of about 50 nucleotides from the 3' end of the 16S rRNA containing the mRNA binding site (Shine-Dalgarno site). Colicinogenic strains are immune to Colicin E3 because they produce protein ImmE3, which binds specifically to colicin E3 and prevents the action. The *col*E3 and *imm*E3 genes are present in a nonself-transmissible plasmid, pColE3-CA38, and they are cotranscribed in an operon.

We have cloned and constitutively expressed *colE3* and *immE3* genes separately. The *imm*E3 gene was integrated and constitutively expressed in the chromosome of different *E. coli* and *Pseudomonas putida* strains. By using PCR techniques and *E. coli imm*E3$^+$ as a recipient strain, the *col*E3 gene was cloned under the control of the Ptac promoter in pVLT31, a promiscuous plasmid able to replicate both in gram-positive and gram-negative bacteria. We have tested the efficiency of our containment system by using different methods of DNA transfer, transformation and conjugation, and different recipient microorganisms representative of the main subgroups of Proteobacteria: *E. coli, P. putida, P. fluorescens, Alcaligenes eutrophus, Comamonas acidovorans, Rhizobium meliloti* and *Agrobacterium tumefaciens*. In all cases we obtained between 10^4-10^6 times fewer plasmid-containing tetracycline-resistant recipients with the plasmid expressing colicin E3 than with the wild type vector.

Our results show that colicin E3 and immunity E3 can be engineered as an efficient gene containment system, and that the bactericidal effect of colicin E3 is manifested in a variety of strains important for environmental purposes. Because the sequence in the 16S rRNA cleaved by colicin E3 is conserved in all prokaryotes, this killing function should be universal in nature, and is thus an excelent potential tool to design barriers that will decrease dispersal of recombinant genes among indigenous microorganisms in ecosystems into which a GEM is deliberately or accidentally introduced.

MAJOR SCIENTIFIC BREAKTHROUGHS

The regulatory circuits involved in the control of the biphenyl pathway of *Pseudomonas* sp. LB400 have been dissected.

A mini-Tn5 cassette bearing the *bph* operon was inserted in the host chromosome of indigenous microbes isolated from different ecological niches, where in situ bioremediation treatment can be done. The recombinant microbes behave like the parental ones in the ecological traits tested.

We have developed the first GEMs specifically designed to degrade pollutants, which were equipped with either circuits for biological containment of the host organism or barriers to limit the rate of lateral transfer of rDNA. This can be considered a scientific breakthrough, as our results confirm that the survival and behavior of GEMs, and rDNA transfer, can be rendered predictable.

COOPERATIVE ACTIVITIES

Lars B. Jensen from the Technical University of Denmark visited J.L. Ramos' laboratory for five months to carry out joint work.
Ute Jakubzik and Birgit Kessler from the GBF visited V. de Lorenzo's laboratory for 4 and 2 weeks respectively, to carry out joint work.
J.L. Ramos visited Dr. Søren Molin and Kenneth N. Timmis to further define collaborative work.
V. de Lorenzo visited all partner laboratories to define and help in the coordination of research activities.
Dr. de Lorenzo and Dr. Ramos acted as external examiners on the Ph.D thesis board for Claus Kristensen's and Lars B. Jensen's dissertations at the Technical University of Denmark.
Dr. D. Dwyer from the GBF visited the University of Cork to plan collaborative research.
Strains were continuously exchanged between all groups, and collaborative links were maintained throughout the project to achieve the proposed goals.
Two plenary meetings, one in Granada and the other in Segovia, were held while the project was running.

PUBLICATIONS

Joint publications

Michán, C., Luming Z., Gallegos, M.T., Timmis, K.N. and Ramos, J.L. (1992). J. Biol. Chem. **267**: 22897-22901.
Hofer, B., Eltis, L.D., Dowling, D.N. and Timmis K.N. (1993). Gene **130**:47-55.
de Lorenzo, V., Cases, I., Herrero, M. and Timmis, K.N. (1993). J. Bacteriol **175**: 6902-6907.
Herrero, M., de Lorenzo, V., Ensley, B. and Timmis, K.N. (1993). Gene **134**: 103-106.
de Lorenzo, V., Fernández, S., Herrero, M. and Timmis, K.N. (1993). Gene **130**: 41-46.
Molin, S., Boe, L., Jensen, L.B., Kristensen, C.S. and Givskov M. (1993). Annu. Rev. Microbiol. **47**: 139-166.
Jensen, J.B., Ramos, J.L., Kaneva, Z. and Molin, S. (1993). App. Environ. Microbiol. **59**: 3713-3717.
O'Sullivan, D.J., Dowling, D.N., de Lorenzo, V. and O'Gara, F. (1994). FEMS Letters, (in press).
Kessler, B., Timmis, K.N. and de Lorenzo, V. (1994). Mol. Gen. Genet. (in press).
Kessler, B., Herrero, M., Timmis, K.N. and de Lorenzo, V. (1994). J. Bacteriol. (in press).
de Lorenzo, V. and Timmis, K.N. (1994). Methods in Enzymology **235**: 386-405.
Díaz, E., Munthali, M., de Lorenzo, V. and Timmis, K.N. (1994). Mol. Microbiol. **13**: 855-861.

Individual Publications

O'Sullivan, D.J. and O'Gara, F. (1992). Microbiological Reviews **56**:662-676.
Fenton, A., Stephens, P.M., Crowley, J., O'Callaghan M. and O'Gara, F. (1992). Appl. Environ. Microbiol. **58**:353-358.
Wagner-Döbler, I., Pipke, R., Timmis, K.N. and Dwyer D.F. (1992). Appl. Environ. Microbiol. **58**:1249-1258.
Dowling, D.N., Boesten, B., O'Sullivan, D.J., Stephen, P., Morris, J. and O'Gara, F. In: Biocontrol Agents: Benefits and Risks. Cambridge University Press. (in press).
Erb, R. and Wagner-Döbler, I. (1993). Appl. Environ. Microbiol. 59.
de Lorenzo, V. and Rojo, F. (1993). In: Avances en Ingeniería Genética. C.S.I.C., Madrid. pp.333-358.
Duque, E., Marqués, S. and Ramos, J.L. (1993). Microb. Rel. **2**:175-177.

Marqués, S. and Ramos, J.L. (1993). Mol. Microbiol. 9:923-929.
Dowling, D.N., Boesten, B., Gill Jr., P.R. and O'Gara, F. (1994). In: Molecular Ecology of Rhizosphere Microorganisms. VCH Weinheim, Germany. (in press).
Dowling, D.N. and O'Gara, F. (1994). In: Current topics in Molecular Genetics. Research Trends, Trivandrum-695 024, India. (in press).
Brazil, D., Callanan, M., Kenefick, L., O'Gara, F., and Dowling, D.N. (1994). (in press).
Brazil, D., Kenefick, L., O'Gara, F. and Dowling, D.N. (submitted)
Pérez-Martín, J., Rojo, F. and de Lorenzo, V. (1994). Microbiological Reviews. **58**: 268-290.
de Lorenzo, V. (1994). Trends in Biotechnology. (in press).
Molin, S. (1993). Curr. Opinion in Biotechnol. 4:299-305.
Givskov, M., Eberl, L., Møller, S., Poulsen, L.K. and Molin, S. (1994). J. Bacteriol. **176**:7-14.
Givskov, M., Eberl, L. and Molin, S. (1994). J. Bact. (in press).

Experimental and modelling studies on the fate in soil of introduced biologically-contained bacteria (BIOT CT-910288)

COORDINATOR:
J.D. van ELSAS, IPO-DLO, Wageningen, NL
PARTICIPANTS:
S. MOLIN, TUD, Lyngby, DK
S. KJELLEBERG, U. Göteborg, Göteborg, S
N. van der HOEVEN, TNO, Delft, NL

BACKGROUND INFORMATION

Successful and effective biological control of plant pests will often require that genetically engineered microorganisms (GEMs) are introduced into the environment *viz* soil. Both the effectiveness and putative hazards of the application require that the fate of the organisms in soil be known and possibly affected, prior to any large-scale field introductions.

OBJECTIVES AND PRIMARY APPROACHES

The project was first directed to the development of strategies for the design of biologically contained, ecologically fit, genetically modified microorganisms (GEMs) and for their sensitive detection in the soil/plant environment. A gene for biological control of insects, *cryIVB, cloned from Bacillus thuringiensis* var morrisoni, has been employed, and inserted into an ecologically adapted carrier strain, *Pseudomonas fluorescens* R2f. Insertion into a second carrier organism, *P. cepacia* P2, was successful, however expression of the gene was poor. The former construct as well as its parent strain and marked derivatives has formed the basis for ecological and ecophysiological studies as well as for the development of strategies to obtain both fitter and selectively disabled derivatives.

Biological containment of strains for release was initially based on the concept of active killing, however mainly due to disappointing results with the functioning of the gef and nuc gene products in *P. fluorescens* R2f, in the course of the project passive containment was conceptually developed as a viable strategy. Since bacteria released into soil will most likely be subjected to starvation, understanding the starvation survival response of released inoculants and tinkering with it was an important objective.

A third objective was the development of a mathematical model which would describe the dynamics of an inoculant when released into soil, taking into account the effects of realistic parameters in soil such as the presence of competing organisms, the soil pore structure in relation to water content and fluxes and the availability of nutrients.

The modified derivative of *P. fluorescens* still has to be tested in a field release study, since (as communicated to the EC in a previous report) the time allotted to the current project has been too small for such an endeavour, given the fact that regulatory hurdles are only now being taken. It is hoped interest can be raised for such a future GEM field endeavour, of which trials will take place in autumn 1994 at IPO-DLO.

RESULTS AND DISCUSSION

A. Development and tracking of soil inoculants for biological control

For biological control of insect larvae in wheat and grass, two rhizosphere-isolated Pseudomonas strains, *P. fluorescens* R2f and *P. cepacia* P2, were selected to serve as carrier strains of the δ-endotoxin gene from *Bacillus thuringiensis* var *morrisoni*, *cry*IVB. Tracking methods for these strains, including selective plating, immunofluorescence and PCR-assisted detection using specific primers and probes, have been developed and successfully applied to monitor the *P. fluorescens* gene delivery vehicles, as well as their heterologous DNA, in soil microcosms (Smalla et al., 1993). Persistence of the gene in the absence of detectable surviving cells was revealed, which indicated the importance of monitoring the inserted gene in addition to cells. The modified R2f strains containing *cry*IVB under the control of P_{tac} were never fitter than the parent strain, however in certain cases showed decreased fitness in soil. We attempted to reduce the putative metabolic load of constitutive gene expression by identifying plant root induced gene expression using promoterless reporter gene inserts and screening for specific responses to plant root exudates. Several such insertion mutants were found and one showed strong and specific reporter gene induction by proline, a component found in exudate of the monocots wheat, maize and grass, but not in that of clover (van Overbeek et al., 1994). Since specific induction in the monocot rhizosphere soil was consistently found, but not in corresponding bulk soil, the promoter was regarded as a good candidate to trigger *cry*IVB gene expression in the monocot rhizosphere. The promoter has been cloned and is currently being subcloned.

We further assessed the usefulness of soil factors such as starvation and low temperature for specific triggering of gene expression in our gene delivery vehicles via random promoterless reporter gene insertions. One mutant each of strains R2f and P2 specifically responding to carbon starvation and one of P2 responding to low temperature were obtained. The promoters identified are useful for insertion of killing principles specifically induced by the relevant soil conditions.

Strategies for containment were based on both passive and active killing, i.e. on interfering with cellular genes involved in survival (passive) as well as on insertion of active killing principles, e.g. the *Escherichia coli gef* gene or *Serratia marcescens* or *Staphylococcus aureus* nuclease *nuc* genes (see further).

B. Passive containment

To develop strategies for passive containment in the oligotrophic (bulk) soil, genes involved in the starvation response were regarded as ideal targets. Therefore, the occurrence of such a response in soil bacteria was analysed. Ideally, containment would be based on the (simplified) concept of bacterial inoculants being mostly metabolically active in the rhizosphere, where nutrient availability is enhanced, and cellular physiology largely shifting to a starvation response in the grossly oligotrophic bulk soil. The work carried out at UG and TUD, together with recent work at IPO-DLO, has now indeed opened avenues for containment based on the starvation response. At UG, the use of two-dimensional gel electrophoresis and direct identification of proteins from 2-D gels by N-terminal amino acid sequencing and functional analysis, of reporter gene fusions and of mutants specifically affected in regulatory functions has greatly advanced knowledge on the starvation response programme in non-differentiating bacteria. Numerous genes and gene products have been identified and their putative functions have been or are being unraveled (see reference list). Specifically, 3 periplasmic space proteins, 3 stress proteins (DnaK, GroEL and Sis1), the major carbon starvation specific protein in

Vibrio sp. DW1, StiI, a responder which presumably regulates the carbon starvation stimulon by RNA regulation, and 10 immediate upshift specific proteins, unique to the transient 3-9 min recovery response upon nutrient addition to starved cells, have been identified. The existence and functionality of a carbon starvation induced stimulon in soil fluorescent pseudomonads as well as the genes and gene products involved has now been shown in work at TUD and IPO-DLO. Passive (as well as active) containment has been proposed to be achievable by tinkering with any of the major, regulatory genes identified. The heat shock protein DnaK was identified as such a major responder, however a *dnaK*$^-$ mutant was also severely impaired in exponential growth, which makes its use as a containment principle exclusively under conditions of carbon starvation doubtful. Further, a spoT double transposon insertion mutant obtained at UG showed severely reduced survival under starvation coupled to a normal growth rate. This gene was shown to be an important regulator in the carbon starvation response, involved in regulating several genes. Recent work at UG has also shown genes specifically involved in the nutrient upshift response, and passive biological containment of organisms in bulk soil might also be based on interference with these.

Passive containment in soil of Pseudomonas gene carrier strains has been tested at IPO-DLO via promoterless reporter transposon insertions into genes specifically induced upon carbon starvation. Mutants selected, the R2f derivative Ra92 and the P2 derivative Pc3, were introduced into medium or soil in a 1:1 ratio with their respective parent strains. In the presence of glucose, there was no difference in growth rates of mutants and parents, whereas in medium without substrate both mutants were seen to be outcompeted by their parent strains. These results clearly showed that both carbon starvation mutants were specifically impaired under carbon limited conditions. Soil experiments were then performed in a model rhizosphere soil system developed by Kuchenbuch, with wheat roots in a root mat exuding into a soil disk containing the inoculant. The soil water regime was kept constant during the experiments. The Ra92/R2f mixture did not show large shifts in the 1:1 ratio applied at the membrane site (model rhizosphere) and in the soil layer between 0 and 2 mm from the membrane, whereas it shifted from 1:1 to about 1:10 one month after inoculation both at 2-8 mm from the membrane and in separate unplanted systems. Hence, the R2f carbon starvation mutant Ra92 represents a potential passively contained strain, since it showed impairment of survival in bulk, but not in rhizosphere soil compared to its parent strain. Secondary Ra92 transposon insertion mutants were produced to check for putative mutations of regulatory genes, which possibly results in severe disruption of the carbon starvation programme; however so far no severely disabled double mutants were found among over 4000 tested.

C. Active containment

For active containment of inoculant strains, several different genes were attempted to obtain differentially induced killing of *P. fluorescens* R2f (TUD). As mentioned in previous reports, neither the *gef* nor the *S. marcescens nuc* gene under the control of constitutive promoters which work in R2f were able to promote adequate killing of this strain. At TUD, recently the extracellular nuclease of Staphylococcus aureus has been cloned, stripped of its export function, and placed on a broad host range suicide introduction system for introduction into a wide range of hosts. Functionality of an active killing principle in soil has been shown with gef induced in the absence of organochloride and repressed in its presence. However, due to pressure favouring non-self-killing mutants, it is doubtful whether the direct approach is fully adequate as a reliable suicide principle. In any case, we propose

the use of the root exudate promoters identified in the *P. fluorescens* gene delivery vehicle at IPO-DLO as powerful regulatory principles for differential regulation of any killer gene in soil (induction) and rhizosphere (repression via a negative loop using the *lac* repressor). To prevent transcription of the suicide gene in the rhizosphere, the *lac* repressor gene *lacI* is then cloned upstream of the promoter whereas the suicide gene controlled by the lac promoter can be used *in trans*. This negative loop construction has been utilized before and is available.

D. Mathematical modelling of bacterial fate

The model predictive of survival and spread of introduced bacteria in soil was based on soil structure affecting bacterial fate. Soil was assumed to contain millions of pores of a few different size classes and an introduced bacterial strain, e.g. a GEM (the invader), was assumed to distribute among the pores of the different size classes. Bacterial movement between pores could only take place with water flow, which also redistributed nutrients. Only one invading bacterial strain could survive per pore, and invasion of already occupied pores was often abortive. Bacteria in larger pores were increasingly subjected to predatory pressure. Bacterial cells were assumed to either grow, die out or become starved and develop hardy (resistant) forms.

The model predicted that generally the GEM invader bacteria ultimately did not become established in soil. Extinction of the invader was, however, often slow, and could take in the order of years. Very harsh conditions in soil were seen to enhance the pore space available for colonization and hence the probability with which the invader was able to expel the indigenous microflora and become established.

Extinction of the invader was slower when it was introduced when soil conditions were becoming more favourable for microbial life than when conditions were worsening.

The model will be employed in the field microplot release study programmed at IPO-DLO in fall 1994.

MAJOR SCIENTIFIC BREAKTHROUGHS

The results lay a firm basis for future successful and safe releases of biocontrol GEMs into soils. The root exudate promoter identified will be used to design strains with plant-induced expression of heterologous genes, representing an ecologically safe and promising introduction strategy. The current knowledge on the starvation response has provided leads to develop better (passive) containment strategies not plagued by selection of mutations, and in particular the *spo*T mutant identified at UG is of major significance. The occurrence of a starvation response programme leading to enhanced cellular resistance in soil pseudomonads, which also is detectable in soil represents a major advance, since it now permits the design of strains disabled in the genes involved. It also provides greater insight in the ecology (activity) of bacterial inoculants in soil/rhizosphere.

The mathematical model developed to our knowledge is the first one which uses key soil features such as soil structure and nutrient and water availability as factors affecting inoculant fate. As such, it is highly valuable, and will be employed in an upcoming field microplot release.

MAJOR COOPERATIVE LINKS

The 4 groups have met twice yearly to discuss progress and research directions. Leo van Overbeek worked for 1 month at the TUD to perform joint experiments. Contacts between TUD and U. Goteborg on the one hand, and IPO-DLO and

TNO on the other hand, have been further frequent in numerous visits and joint endeavours.

PUBLICATIONS

Joint publications

Molin, S. and Kjelleberg, S. 1993. Release of Engineered Microorganisms: Biological Containment and Improved Predictability for Risk Assessment. AMBIO **22**: 242-245.

Van der Hoeven, N. and van Elsas, J.D. 1994. A Model based on Soil Structural Aspects Describing the Fate of Genetically Modified Bacteria in Soil. Ecol. Modell. (submitted).

Van Overbeek, L.S., Eberl, L., Givskov, M., Molin, S. and van Elsas, J.D. 1994. Stress Induced Resistance in *Pseudomonas fluorescens* Introduced into Soil. Applied and Environmental Microbiology (submitted).

Individual publications

Molin, S., Jensen, L.B., Kristensen, C.S. and Ramos, J.L. 1993. Programmed Survival of Soil Bacteria. In: Trends in Microbial Ecology (eds. Guerrero and Pedros-Alio), 175-178.

Molin, S., Boe, L., Jensen, L.B., Kristensen, C.S., Givskov, M., Ramos, J.L. and Bej, A.K. 1993. Suicidal Genetic Elements and their Use in Biological Containment of Bacteria. Ann. Rev. Microbiol. **47**: 139-166.

Molin, S. 1993. Environmental Potential of Suicide Genes. Curr. Opin. Biotechnol. **4**: 299-305.

Molin, S. 1993. Designing Microbes for Release into the Environment. Sci. Progress **76**: 1-13.

Givskov, M., Eberl, L., Moller, S., Poulsen, L.K. and Molin, S. 1994. Analysis of General Cross-Protection, Cell Morphology, and Macromolecular Content. J. Bacteriol. **176**: 7-14.

Givskov, M., Eberl, L. and Molin, S. 1994. Responses to Nutrient Starvation in Pseudomonas putida strain KT2442: Two-dimensional electrophoretic analysis of starvation and stress inducible proteins. J. Bacteriol. (in press).

Smalla, K., Cresswell, N., Mendonca-Hagler, L.C., Wolters, A. and van Elsas, J.D. 1993. Rapid DNA extraction protocol from soil for polymerase chain reaction-mediated amplification. J. Appl. Bacteriol. **74**: 78-85.

Van Elsas, J.D. and Van Overbeek, L.S. 1993. Bacterial Responses to Soil Stimuli. In: Starvation in Bacteria, Kjelleberg, S. (Ef.), pp. 55-79. Plenum Press, New York.

Van Overbeek, L.S. and van Elsas, J.D. 1994. Root exudate induced promoter activity in *Pseudomonas fluorescens* mutants in the rhizosphere of wheat. Appl. Environ. Microbiol. (in press).

Van Overbeek, L.S. and van Elsas, J.D. 1994. Use of environmentally-triggered promoters for specific differential gene expression in bacterial inoculants in soil and rhizosphere. In: M.H. Ryder, P.M. Stephens and G.D. Bowen (Eds.), Improving Plant Productivity with Rhizosphere Bacteria, CSIRO Div. of Soils, Australia, pp. 238-244.

Kjelleberg, S., Albertson, A., Flärdh, K., Holmquist, L., Jouper-Jaan, A., Marouga, R., Östling, J., Svenblad, B. and Weichart, D. 1993. How do non-differentiating bacteria adapt to starvation? Antonie van Leeuwenhoek **63**: 333-341.

Flärdh, K., Cohen, P.S. and Kjelleberg, S. 1992. Ribosomes exist in large excess over the apparent demand for protein synthesis during carbon starvation in marine *Vibrio* sp. strain CCUG 15956. J. Bacteriol. **174**: 6780-6788.

Flärdh, K. and Kjelleberg, S. 1994. Glucose upshift of long-term carbon-starved marine *Vibrio* strain sp. S14 causes severe amino acid starvation and induction of the stringent response. (submitted).

Holmquist, L., Jouper-Jaan, A., Weichart, D., Nelson, D.R. and Kjelleberg, S. 1993. The Induction of Stress Proteins in three Marine *Vibrio* during Carbon Starvation. FEMS Microbiol. Ecol. **12**: 185-195.

Holmquist, L. and Kjelleberg, S. 1993. The Carbon Starvation Stimulon in the Marine *Vibrio* sp. S14 (CCUG 15956) includes three periplasmic space protein responders. J. Gen. Microbiol. **39**: 209-215.

Jouper-Jaan, A., Goodman, A.E. and Kjelleberg, S. 1992. Bacteria Starved for Prolonged Periods Develop Increased Protection against Lethal Temperatures. FEMS Microbiol. Ecol. **101**: 229-236.

Östling, J., Holmquist, L., Flärdh, K., Svenblad, B., Jouper-Jaan, A. and Kjelleberg, S. 1993. Starvation and Recovery of *Vibrio*. In: Starvation in Bacteria, Kjelleberg, S. (Ed.), pp. 103-127. Plenum Press, New York.

An experimental approach to investigate horizontal gene transfer between organisms (BIOT CT-910287)

COORDINATOR:
R.P. OLIVER, UEA, Norwich, GB
PARTICIPANTS:
M.-A. GRANDBASTIEN, INRA, Versailles, F

BACKGROUND INFORMATION

Horizontal gene transfer is the movement of genetic information between species. Although there is a great deal of evidence that horizontal transfer occurs between bacteria and in some specialised situations between bacteria and plants, there is no direct evidence of transfer between eukaryotic organisms (plants, animals and fungi). There is, however a great deal of circumstantial evidence of transfer between eukaryotes. The aim of this project was to directly assay for transfer between plants and fungal pathogens. Such a discovery would have great impact on the biological species concept and on considerartions of the safety of release of genetically-modified organisms.

Mobile genetic elements have been characterized in a wide range of organisms, from bacteria to higher eukaryotes. One major type of mobile element is the retrotransposon family, whose members are structurally and functionally related to animal retroviruses, and which replicate through an RNA intermediate. They represent the major class of transposable elements in animal genomes, and have been shown to be also ubiquitous components of plant and fungal genomes. Retrotransposons provide good examples of sequences from unrelated organisms showing unexpectedly high degrees of sequences homology; their pattern of similarity is highly suggestive of horizontal gene transfer. Moreover their replication cycle, involving cytoplasmic virus-like particles, provides an obvious mechanism for horizontal transfer.

Both these considerations indicate that retrotransposons are very likely to have transferred and to be capable of transfer between organisms and were therefore the subject of this study.

OBJECTIVES AND PRIMARY APPROACHES

The ultimate goal of this project was to experimentally measure the possibility of horizontal transfer of retrotransposons between a plant and its biotrophic fungal pathogen. The chosen model involved the fungal pathogen *Cladosporium fulvum*, and its host species, tomato, and the Tnt1 and CfT-1 elements, isolated respectively from tobacco and from *C. fulvum*.

The experimental approach involved the creation of chimaeric retrotransposons, containing marker genes, such as antibiotic resistance, whose transfer from the plant to the fungus can be detected. In the next step, these chimaeric elements were to be introduced in the tomato genome. The marked tomato strains are then to be infected by *C. fulvum*, and fungal spores acquiring the marker gene sought. The laboratory of Dr R.P. Oliver, at UEA in Norwich, is working on the tomato-C.fulvum interaction, has isolated the fungal CfT-1 element, and provides the fungal technology needed for this project, while the laboratory of Dr. M.A.

Grandbastien, at the INRA in Versailles, has isolated the tobacco Tnt1 element, and provides the plant technology needed for tomato transformation experiments.

In addition to this major goal, the project involved other preliminary studies, important to establish the conditions in which horizontal transfer was might take place.
In Norwich, we have investigated expression of the CfT-1 element and shown that it is likely to be expressed during infection. We have also tested various components of the chimaeric constructs prepared by M A Grandbastien for operation in *C. fulvum*. In addition we have established the phylogeny of *C. fulvum* and fungal retrotransposons by comparison of sequences.
In Versailles, these preliminary studies involved the assessment of the expression of the tobacco Tnt1 element in the tomato genome, and also the conditions in which the Tnt1 element was expressed in the plant. This was an important prerequisite for this work, since nothing was known about the expression of plant retrotransposons, whose discovery and analysis is a brand new field.

RESULTS AND DISCUSSION

1) Norwich

1.1) Construction of chimaeric CfT-1 elements

Chimaeric constructs in which plant and fungal selectable antibiotic resistance genes have been inserted in the fungal retrotransposon have been constructed. These are being transformed into *C. fulvum* and have been sent to Versailles for transformation into tomato.

1.2) Expression of CfT-1 during infection

Expression of CfT-1 in *C. fulvum* was tested by analysis of fusions between the promoter of the transposon (the LTR), and a reporter gene, GUS. Transformation of this construct into *C. fulvum* has demonstrated that expression of the LTR is enhanced during starvation, a condition known to prevail during infection. We have also demonstrated that the plant 35S promoter, used to drive antibiotic resistance, operates in the fungus.

In addition we have undertaken various studies to examine the possible causes of the unexpected level of sequence homology between retrotransposons. In doing so we have established the phylogeny of *C. fulvum* in relation to Ascomycete fungi and other members of the *Cladosporium genus*, using rapid molecular methods. The sequences have been deposited in the databases. A retrotransposon has been found in *Cladosporium cladosporioides*, called CcT-1. Phylogenetic analysis of CcT-1, CfT-1 and other retrotransposons was found to be consistent with vertical transfer, but this does not rule out horizontal transfer.

2) Versailles

2.1) Tnt1 expression in tomato

Previous results had shown that Tnt1 was poorly expressed in the tobacco plant, excepted in roots, but that Tnt1 expression was highly induced during tobacco leaf mesophyll protoplast isolation, due to the application of fungal extracts (*Trichoderma viride*) containing cell wall hydrolases.

Tnt1 expression was further analysed, both in tobacco and in tomato, by creating transgenic plants containing a transcriptional fusion between the Tnt1 LTR, carrying the promoter and regulatory sequences, and the GUS reporter gene (LTR-GUS construct), and by analyzing the GUS activity of these plants. These studies

showed that Tnt1 expression is highly induced in tobacco by several factors of microbial origin, such as fungal elicitins produced by *Phythophtora* sp, and culture supernatants of the bacterium Erwinia chrysanthemi. All the inducing factors tested have in common the ability to elicit in tobacco plant defence responses of the hypersensitive type, and Tnt1 activation is tightly correlated to the necrotic activity of elicitins. In tomato, Tnt1 expression is also highly induced by *Trichoderma viride* fungal extracts and by *Phythophtora* fungal elicitins. Moreover, a recent analysis has shown that a strong induction of Tnt1 expression is also found with *C. fulvum* culture extracts, containing non-specific elicitors of necrosis.

All inducing factors tested have in common the ability to elicit plant necrotic defence responses leading to a non-specific incompatible interaction. It is therefore necessary to determine if Tnt1 induction is restricted to incompatible interactions, and the gene for gene interactions existing between tomato and *C. fulvum* provides an excellent model for this study. In this system, an interaction between the product of specific elicitors produced by *C. fulvum* avirulence genes (*Avr*) and the corresponding plant resistance genes (Cf) is responsible for a plant necrotic response leading to resistance to the pathogen. We have therefore introduced, by genetic crosses, the Cf9 resistance gene into our LTR-GUS transgenic tomato lines, and tested the effect of the *C. fulvum* Avr9 gene product. Our latest results indicate that the Avr9 gene product activates Tnt1 expression in tomato lines containing the Cf9 resistance gene, but not in those devoid of the resistance gene

These results show that the tobacco Tnt1 retrotransposon is expressed in tomato, and represents therefore a suitable tool for the experimental testing of horizontal transfer between tomato and *C. fulvum*. Tnt1 expression in tomato, as in tobacco, is linked to the plant response to microorganisms, but our results suggest that this expression might be higher in situation leading to resistance of the plant to the pathogen. Further experiments, such as real infections with virulent and avirulent *C. fulvum* races, are now underway to determine if Tnt1 induction is a specific response to the presence of the avirulence gene product, or if other compounds produced during this plant-pathogen interaction are also able to induce Tnt1 and to modulate these observations.

Our results show however that modifications of the Tnt1 promoter, to create constitutively expressed elements, are probably required to ensure a high level of induction during a compatible interaction.

2.2) Construction of chimaeric Tnt1 elements

Chimaeric Tnt1 elements containing marker genes that can be selected for in the fungal genome, such as the gene coding for hygromycin resistance, have been constructed. Marker genes have been inserted at the end of the element's ORF, after introduction of an unique restriction site, and two types of elements have been constructed:

- a Tnt1 element carrying the hygromycin resistance under control of the 35S CaMV promoter, expressed both in tomato and *C. fulvum*
- a similarly marked, but constitutively expressed, Tnt1 element. In this element, the 5' LTR sequences upstream from the major transcriptional start (U3 region) have been replaced by the 35S CaMV promoter enhancer sequences. Previous results have shown that this substitution, which should not affect the element's transpositional activity, induces a high level of constitutive expression of the LTR in plant leaves.

In parallel, chimaeric CfT-1 elements have been constructed in Norwich, and sent to us. The chimaeric CfT-1::HYGRO and Tnt1::HYGRO elements have then been cloned into the T-DNA region of the pBIN19 binary vector, and these constructs have been introduced into *Agrobacterium tumefaciens* strain C58C1.

2.3) Tomato transformation

Tomato cotyledons pieces have been infected by these *A. tumefaciens* clones. Transformation events, corresponding to the transfer of the T-DNA region to the plant genome, have been obtained through the selection of calli resistant to kanamycin carried on the T-DNA region. The regeneration of tomato plantlets from resistant calli is expected soon, and will generate transgenic tomato plants containing the chimaeric retrotransposons.

Transfer of the elements to the fungus will then be assayed at the UEA, Norwich, UK, by infection of these transgenic tomatoes with *C. fulvum*, and selection of resulting spores for resistance to hygromycin.

PUBLICATIONS

Individual publications

M. Dewey, A. Schots & R.P. Oliver eds. (1994). Modern Assays for Plant Pathogenic Fungi; Identification, Detection and Quantification. CABI, Wallingford.

R.P. Oliver (1992). A model system for the study of plant-fungal interactions: Tomato leaf mould caused by *Cladosporium fulvum*. In: DPS Verma ed. *Molecular Signals in Plant-Microbe Communication.* CRC Press. pp. 297-308.

M.J. McPherson, R. Oliver & S.J. Gurr (1992). The polymerase chain reaction. In: *Molecular Plant Pathology: A Practical Approach.* Eds Gurr SJ, McPherson MJ and Bowles DJ, IRL pp. 123-146

M.T. McHale, I.N. Roberts, S.N. Noble, C. Beaumont, M.P. Whitebread, D. Seth & R.P. Oliver (1992). CfT-1, An LTR-Retrotransposon in *Cladosporium fulvum*, a fungal pathogen of tomato. *Molecular and General Genetics,* **233**: 337-347

R.P. Oliver (1993). The use of nucleic acid-based methods in the detection and identification of plant pathogens. In: *Principles of Diagnostic Methods in Plant Pathology.* Ed RTV Fox. CAB International.

R.P. Oliver, M.L. Farman, J.D.G. Jones & K.E. Hammond-Kusack KE (1993.) Use of fungal transformants expressing β-glucuronidase activity to detect and measure hyphal biomass in infected plant tissues. *Molecular Plant Microbe.* (in press).

M.J. Coleman, M.T. McHale, J. Arnau, A. Watson & R.P. Oliver (1993). Cloning and characterisation of telomeric DNA from *Cladosporium fulvum*. Gene **132**: 67-73

M.D. Curtis J. Gore & R.P. Oliver (1994). The phylogeny of the tomato leaf mould fungus, *Cladosporium fulvum* syn. *Fulvia fulva* by analysis of rDNA sequences. Current Genetics, In press

I.N. Roberts, D.C. Glayzer, D.J. Jeenes, R.P. Oliver & D.B. Archer (1992). Retrotransposon-related sequences in the genome of *Aspergillus niger*. Abstract at International Conferences on Computing in Molecular Biology (Genes, Proteins and the Computer)

D. Gmayzer D, I.N. Roberts, M. Whitehead, R.P. Oliver & D. Archer (1993). In search of retrotransposons in *Aspergillus niger*. Abstract presented in the Fungal Genetics Congress, Asilomar, California

S. Pouteau, M.A. Grandbastien, M. Boccara (1994). Microbial elicitors of plant defense responses activate transcription of a retrotransposon. Plant Journal. **5**: 535-542

M.A. Grandbastien (1992). Retroelements in higher-plants. Trends In Genetics **8**: 103-108

Safety of genetically engineered retroviruses used for gene transfer (BIOT CT-910286)

COORDINATOR:
F.S. PEDERSEN, Dept. of Molecular Biology, Aarhus Univ., Aarhus, DK
PARTICIPANTS:
W.H. GÜNSBURG, GSF, Oberschleissheim, D
J. GUARDIOLA, IIGB, Napoli, I
G. VERDIER, Université Claude Bernard Lyon I, Lyon, F
C.-M. CALBERG-BACQ, Université de Liège, Liège, B

BACKGROUND INFORMATION

Retroviral vectors (RVs) are presently used for the transfer of genes for research, industrial and clinical uses for genetic therapy of various cancers, genetic diseases and acquired immunodeficiency. They offer many advantages as gene transfer vehicles including efficient delivery and integration of the delivered genes into the host cell DNA. However, a number of safety concerns needed to be addressed with respect to using such vector systems. The major problem associated with the use of RVs is the generation of infectious virus that can be indefinitely transmitted. This is undesirable because retroviruses are associated with tumorigenesis and immunopathogenesis. Currently used RVs deliver and express genes in many cell types. RVs that deliver and/or express genes to/in predefined cell types would be safer and possibly more efficient.

OBJECTIVES AND PRIMARY APPROACHES

To assess and improve the safety of retroviral mediated gene transfer. The major effort was to determine the frequency and mechanisms of recombination with a view to constructing safer RVs that are less able to recombine. Additionally features such as

(1) physical stability of virus particles,
(2) stability of retroviral sequences once integrated in the host cell DNA,
(3) stability and fidelity of gene expression and the influence of environmental stimuli and
(4) the influence of the site of integration upon gene expression were investigated.

A second effort was to design and test new RV systems with the aim of improving safety and specificity. Specific parts of the work attempt to control the reverse transcription step as a means to further control vector mobilization, and the provirus integration step in terms of targeting to predefined sites in host DNA. Strategies have also been developed for the construction of RVs that are targeted, either at the level of the infection event or at the level of expression.

RESULTS AND DISCUSSION

1) Determination of the risk of *in vivo* gene transfer in non-target cells following different routes of entry (UL). Double expression ecotropic murine leukaemia virus (MLV) vectors were produced with a selection gene (*neo*) and either a *lac*Z marker gene or a potent oncogene. In vivo gene transfer was easily demonstrated by X-gal staining of cells on frozen sections or by PCR amplification of the neo sequence in genomic DNA. Marker gene expression was detected 2.5 months after

injection of the lowest virus amount assayed (2×10^4). Efficiency was increased using methods to improve the viral infectious titre of the injected preparation or by injection of virus-producing cells instead of free viruses. In spite of their very short half-life in the organism, RVs injected locally, frequently led to gene transfer but the integrated foreign gene was not always expressed. The viral spreading was however very limited after injection into mammary tumours. No harmful effect could be observed when a potent oncogene (*hst*) was introduced by different ways into rapidly developing new-born mice, which represents a worst case condition. These results demonstrate the safety of the local *in vivo* use of RV preparations that are not contaminated with replication-competent viruses.

2) Assessment of male germinal line infection by retroviral vectors (GSF, IIGB). A system has been developed to test the risk of infection by retroviral particles of spermatogonia in mice; this methodology involves transplantation of retrovirus producing cells into the testes of the animals. Our results indicate that even in the worst-case conditions transfer of the tracer RV genome to the germ-line of the offspring is extremely rare.

3) Assessment of the stability of retroviral vector particles and proviruses and the influence of environmental stimuli (UL, UA). We set up conditions to concentrate vector particles and studied their stability after freezing at −70°C and thawing at 37°C in either mouse blood or culture medium and the functional half-life under '*in vivo*' conditions was determined as 2.6 h (UL). The stability of expression of 21 single gene vector proviruses available as single non-selected integrants in cell clones has been monitored over a five months cultivation period (UA). Although all proviruses are physically stable, major differences in their long-term stabilities of expression were observed. The steroid hormone dexamethasone was found to stimulate the expression from already expressing proviruses in a systematic manner but had no effect on completely inactive proviruses.

4) Development of a set of avian retroviral packaging cells with distinct receptor specificities (UCB). Avian leukosis virus (ALV) based vector systems and safe packaging cell lines that produce high titer helper-free preparations of these vectors were developed. The safest packaging cell line has been constructed by separating the *gag-pol* and *env* genes on two transcriptional units, deleting the packaging sequences and the 3'non-coding regions including the 3'LTR. Moreover, genes for selectable markers have been inserted into the transcriptional units of viral genes allowing us to produce the packaging functions under selective pressure. Use of two separate plasmids has allowed the generation of packaging cells with subgroups A,B,C and E envelope specificities from cell line producing ALV Gag and Pol proteins into which env genes from different viruses were introduced.

The packaging cells generated from quail cells have been checked for absence of production of replication-competent viruses and for stable production of retroviral vectors in high titres. In contrast, the use of a chicken cell line as a helper that bears and expresses ALV related sequences resulted in the generation of replication-competent viruses after various times. Some endogenous RNAs were found to be packaged into virus particles. Superinfections of packaging cells by a *lac*Z retroviral vector after co-cultivation of packaging cells of two subgroup specificities (A and E) as already described for mammalian vector systems ('ping-pong') did not result in increased retroviral vector titers but caused the emergence of replication-competent viruses which result from recombination between the viral sequences of the packaging cell lines and the cis-acting sequences of the *lac*Z retroviral vector. Replication competent viruses were shown to arise only in co-cultures of

vector-producing packaging cells with two host ranges where multiple replication cycles are possible. However, recombinant viruses not competent for replication have been detected and characterised in vector-producing packaging cells cultivated alone. The data have led to the choice of appropriate protocols that minimize the risk of emergence of unexpectable recombinant viral forms.

5) **Steps towards redirecting the host-range of retroviral vectors (UCB).** A second approach to targeting involves the design of new viral receptor specificites by modifying retroviral *env* genes. On the basis of comparative studies of various ALV SU envelope proteins and theoretical structure considerations four small regions have been identified as potential binding domains to the specific receptors. Several constructs have been made in order to substitute these domains by peptides that specifically interact with known target cell membrane proteins as an attempt to generate RVs with predefined cell tropisms. For example, ALV based RVs have been produced with a modified avian ecotropic subgroup A *env* gene, by replacing the expected binding domains by a 16 amino acid long peptide able to recognize integrins exposed on the surface of some mammalian cells. Coding sequences of other peptides (such as epidermal or transforming growth factors) have been used to produce chimaeric *env* genes that have been tested with respect to production of modified Env proteins, viral production, infection efficiency and redirection of the host range. Some SU mutants have been found able to infect, with low efficiency, mammalian cells expressing the EGF receptor.

6) **Steps towards the development of transcriptionally targeted retroviral vectors (GSF, IIGB).** The GSF group aims to develop RVs that are targeted towards the mammary gland. A negative regulatory element in the whey acidic protein promoter, identified as a potential mammary specific determinant has been introduced into a MLV based retroviral vector carrying a marker gene in place of the viral promoter. Recombinant virions have been used to infect cells *in vivo* and in vitro and expression of the marker gene was detected only in mammary cells. Construction of potentially mammary specific RVs based on mouse mammary tumour virus (MMTV) revealed two virally encoded regulatory factors, a transcriptional repressor (Naf) and a superantigen (Sag) which may stimulate, delete or energise whole classes of T cells. Naf and Sag were shown to be independent MMTV products using a facile assay we have developed. A novel MMTV promoter has been identified upstream of the previously described promoter that gives rise to transcripts encoding Sag and Naf. These factors may prove useful as a means to manipulate or target expression of RVs.
The fidelity and efficiency of expression of a tissue-specific promoter cloned in a retroviral vector was analyzed by the IIGB group using the MHC class II HLA-DQA1 promoter known to be expressed in cells of the lymphoerythropoietic system. The structure of the HLA-DQA1 promoter was analyzed by deletion mapping and CAT assays. DNA segments encompassing the regulatory signals were fused to the reporter gene *hph*, conferring hygromycin B resistance, and introduced into a suitable retroviral vector, pLJ. The promoter-distal and -proximal HLA-DQA1 controlling regions were thus identified. In conventional vectors the expression of the reporter gene was correctly regulated, whereas it was below detection in pLJ. The mechanisms causing this phenomenon were studied. We found that lack of expression of a cellular promoter cloned in a RV is due to the interference of sequences present both in the vector and in the promoter. The activity of the cellular promoter could be restored by inactivating some of these sequences by mutation. We are now in the process of identifying those sequences.

7/ The influence of the site of integration on vector proviruses (UA). The contributions of proviral enhancer strength and the site of integration to the level and stability of expression of vector proviruses have been studied. We employed two isogenic MLV derived vectors differing in the strength of the LTR-enhancer, the strong enhancer being derived from the high-leukaemogenic SL3-3 virus and the weak enhancer from the low-leukaemogenic Akv virus. Cell clones with single integrated proviruses were isolated without selection for vector expression. Our results point to an effect of the integration site on the level as well as on the stability of expression. These two effects show no pattern of correlation. To identify the site of integration a simple hemi-specific PCR method was developed. With this method that allows the direct determination of the host sequences flanking an integrated provirus we have begun an analysis of the transcriptional activity and the chromatin structure at pre-integration sites to gain insight into the biochemical basis for the observed chromosomal position effects. Our results demonstrate that a vector carrying the weak Akv transcriptional enhancer in the LTR may give efficient and stable expression levels. This enhancer is less active in terms of insertional activation of host genes than that of the commonly used highly pathogenic Moloney MLV. These results may therefore point to possibilities for the development of RVs with improved safety features.

8) The specificity of tRNA primer usage for murine leukaemia virus based vectors (UA). All retroviruses and hence retroviral vectors use a specific host cell derived tRNA primer for reverse transcription. The tRNA primer anneals to a specific sequence, the primer binding site, located near the 5'end of the viral genome. We have studied the specificity of tRNA primer usage and of primer-binding site interactions as a possible point for control of mobilization of murine leukaemia virus derived vectors. These vectors normally use a primer tRNA specific for proline. Our results show that vectors carrying mutated primer binding sites that match other tRNA species (acceptors for glutamine, lysine and methionine) can replicate efficiently by using these tRNA species. Additional mutational studies demonstrate that the replication machinery may tolerate several mismatches between the tRNA primer and the primer binding site. Vectors carrying a primer binding site sequence that does not match any naturally occurring tRNA species are severely impaired in their replication capacity (about 10^5 fold reduction in titre). Rare transduction events have been shown to result from either recombination with a specific family of endogenous murine retroviruses (employing tRNA-gln primers) or from erroneous replication using a tRNA-pro primer. On basis of these results we have designed an artificial tRNA-like primer and shown that packaging cells carrying this RNA allow efficient vector transfer. These results therefore point to a new strategy for biological containment of retroviral vectors based upon the use of artificial primers for reverse transcriptase.

9) Alternative pathways for proviral integration (GSF, UA, UCB). To study rare recombination events between exogenous and endogenous proviruses and to attempt the development of locus-targeted vectors, we have developed transfer systems that are impaired in the normal integration machinery. Helper free MLV based packaging cell lines yielding virus particles, albeit at low titres, defective in the integrase have been produced (UA). The GSF group has developed murine leukaemia virus based RVs designed for locus-targeting to the herpes simplex virus thymidine kinase gene by homologous recombination. The UCB group has used vectors provided with an additional internally located integration sequence. The results demonstrate that vectors endowed with such an additional integration sequence lead to disorganized proviral integration that may allow specific expression of an internal transcription unit. The UA group has begun to apply the

sequence specific Cre-lox recombination system derived from bacteriophage P1 to control recombination events (integration, excision, inversion etc.) involving RVs.

MAJOR SCIENTIFIC BREAKTHROUGHS

The discovery of a second promoter in the long terminal repeat of mouse mammary tumour viruses represent a significant, unexpected scientific result; this finding may open a new venue for the design of retroviral vectors. The development and characterization of a set of avian retroviral packaging cells with distinct receptor specificities is of major technical interest as is the genetic modification of the binding domain of avian retrovirus envelopes as a step towards redirection of the host-range of retroviral vectors. The discovery of the versatility of tRNA primer usage for murine leukaemia virus based vectors and of the flexibility in tRNA primer binding site interactions is of mainly scientific interest, while the further demonstration that retroviral vectors may be designed to depend upon genetically designed primers is of technical interest for further exploitation.

MAJOR COOPERATIVE LINKS

Four half-yearly meetings of two to three days duration with all groups and selected guest speakers have been held to review data and plan further experiments. The meetings have allowed detailed comparisons of the variety of biological models employed, i.e. avian viruses (UCB), murine leukaemia viruses (UA, GSF, IIGB, UL), mouse mammary tumour viruses (GSF, UL), haematopoietic targeting (IIGB), mammary gland targeting (GSF, UL) germ line integration (GSF, IIGB, UCB). Mutual exchanges of vectors, cell lines, reagents etc. have taken place and a number of collaborative studies have been initiated, in particular between GSF and IIGB on germ line integration, between GSF and UL on mammary gland targeting, between UA and GSF on basic vector design and transcriptional control, each resulting in short interlaboratory visits and in exchange of vectors, cell lines, reagents etc. Retroviral assay systems and virus purification have been standardized between the GSF and UL groups.

PUBLICATIONS

Individual publications

Allioli, N., Legras, C., Thomas, J.L., Chebloune, Y., Nigon, V.M. and Verdier, G. (1994). Use of retroviral vectors to introduce and express the beta-galactosidase marker gene in cultured chicken primordial germ cells (accepted with modifications in 'Developmental Biology').

Cosset, F.L., Girod, A., Flamant, F., Drynda, A., Ronfort, C., Valsesia, S., Molina, R.M., Faure, C., Nigon, V.M. and Verdier, G. (1993). Use of helper cells with two host ranges to generate high-titer retroviral vectors. Virology **193**, 385-395.

Cosset, F.L., Ronfort C., Molina, R.M., Flamant, F., Drynda, A., Benchaibi, M., Valsesia, S., Nigon, V.M. and Verdier G.. (1992). Packaging cells for Avian leukosis virus (ALV)-based vectors with various host ranges. J. Virol. **66**, 5671-5676.

Del Pozzo, G., Ciullo, M., Autiero, M., and Guardiola, J. (1994). Control of nucleocytoplasmic HLA-DRA mRNA partitioning by interaction of a retention signal with compartmentalized proteins. J. Mol. Biol., **240**, 193-204.

Duch, M., Paludan, K., Lovmand, J., Pedersen, L., Jørgensen, P. and Pedersen, F.S. (1993). A correlation between dexamethasone inducibility and basal expression levels of retroviral vector proviruses. Nucleic Acids Res. **21**, 4777-4782.

Günzburg, W.H., Heinemann, F., Wintersperger, S., Miethke, T., Wagner, H., Erfle, V. and Salmons, B. (1993). Endogenous superantigen expression controlled by a novel promoter located in the MMTV long terminal repeat. Nature **364**, 154-158.

Günzburg, W.H., Saller, R. and Salmons, B. (1995). Retroviral vectors directed to predefined cell types for gene therapy. Biologicals, (in press)

Günzburg, W.H. and Salmons, B. (1992). Factors controlling the expression of mouse mammary tumour virus. Biochem. J. **283**, 625-632.
Janka, I., Günzburg, W.H., Brem, G., Erfle, V. and Salmons, B. (1993). Sub-retroviral particles as gene transfer vectors that by-pass retrovirus-receptor interaction restrictions. Biochem. Mol. Biol. Int. **31**, 349-357.
Kolb, A.F., Günzburg, W.H., Albang, R., Brem, G., Erfle, V. and Salmons, B. (1994) A proximal regulatory element in the mammary specific whey acidic protein promoter. J. Cell. Biochem. **56**, 245-261.
Lund, A.H., Duch, M., Lovmand, J., Jørgensen, P. and Pedersen, F.S. (1993). Mutated primer binding sites interacting with different tRNAs allow efficient murine leukemia virus replication. J. Virol. **67**, 7125-7130.
Ombra, M.N., Autiero, M., De Lerma-Barbaro, A., Baretta, R., Del Pozzo, G., Guardiola, J. (1993). Recognition of distinct *HLA-DQA1* promoter elements by a single nuclear factor containing *Jun* and *Fos* or antigenically related proteins. Nucleic Acids Res. **21**, 1811-1818.
Salmons, B. and Günzburg, W.H. (1993) Targeting of retroviral vectors for gene therapy. Human Gene Therapy **4**, 129-141.
Sørensen, A.B., Duch, M., Jørgensen, P. and Pedersen, F.S. (1993). Amplification and sequence analysis of DNA flanking integrated proviruses by a simple two-step polymerase chain reaction method. J. Virol. **67**, 7118-7124.
Sørensen, M.S., Duch, M., Paludan, K., Jørgensen, P., and Pedersen, F.S. (1992). Measurement of hygromycin B phosphotransferase activity in mammalian cell extracts by a simple dot-assay. Gene **112**, 257-260.
Toscano, M., Grignet, C., Hajitou, A. and Calberg-Bacq, C.-M. (1994). Use of retroviral vectors for *in vivo* gene transfer into mouse mammary tissues. Arch. Int. Physiol. Biochim. Biophys., (in press).
Valsesia-Wittmann S., Drynda, A., Deleage, G., Aumailley, M., Heard, J.M., Danos, O., Cosset, F.L. and Verdier, G. (1994). Modification in the binding domain of avian retrovirus envelope protein to redirect the host-range of retroviral vectors. J. Virol. **68**, 4609-19.
Wintersperger, S., Indraccolo, S., Miethke, T., Günzburg, W.H. and Salmons, B. (1994) A transient assay for gene expression studies in B-lymphocytes and its use for superantigen assays. BioTechniques **16**, 882-886.

In addition 4 manuscripts with support from the contract have been submitted and several manuscripts with support from the contract are in preparation.

Patent proposal
A.H. Lund, M. Duch, J. Lovmand, P. Jørgensen and F.S. Pedersen 'Retroviral vector propagation system dependent upon artificial primers for reverse transcriptase' under finalization with support from C.E.C. DG XIII.

Assessment of environmental impact from the use of live recombinant virus vaccines (BIOT CT-910289)

COORDINATOR:
(1) D.N. BLACK, I.A.H., Pirbright Lab, Pirbright, GB
PARTICIPANTS:
(2) A.L.J. GIELKENS, CDI., Lelystad, NL
(3) M.A. SKINNER, IAH, Compton Lab., Compton, GB
(4) P.P. PASTORET, Univ. de Liège, FMV.IVI., B
(5) R.M. GASKELL, Univ. of Liverpool, Neston, GB

BACKGROUND INFORMATION

The development of recombinant virus vaccines and their subsequent use will require the introduction of regulatory policies as yet largely undefined. Such policies must be based upon a wide and sound scientific data base. The programme is aimed at assessing the biological characteristics of poxvirus recombinant vaccines and assessing the environmental impact (if any) from the use of such live recombinant virus vaccines.

OBJECTIVES AND PRIMARY APPROACHES

Recombinant viruses containing genes from heterologous viruses will be constructed and examined for their ability to grow in target and non-target animals and cell lines. Using model recombinant viruses a quantitative assessment will be made of the rate of possible recombination between these viruses.

RESULTS AND DISCUSSION

1) Pirbright

Recombinant capripoxviruses containing the F or H genes of RPV have been constructed and tested. These results show that they are efficacious in the protection of cattle, sheep and goats against challenge with capripox and morbilli viruses. Analysis of the purified recombinant viruses showed that the virus particles contain the H protein and possibly the F protein.The results with the F protein are equivocal as the F protein does not give clear specific reactions. The parent capripoxvirus grows only in cells derived form the natural hosts (cattle, sheep and goats) the most commonly used cells being primary/secondary lamb testis cells. The recombinant viruses grew readily in the lamb testis cells. In cell lines not normally permissive to capripoxvirus growth (Hela, BHK and Vero cells) some early stages of replication of the recombianant viruses could be detected. However, they did not appear to undergo a complete replication cycle and progeny virus were not obtained. The transfer vector for construction of the capripoxvirus containing the PRV gII gene has been constructed and the transfection/recombination experiment will be started shortly so that we can provide the capripoxvirus-gII recombinant for testing at Lelystad. These results would indicate that the recombinant capripoxviruses have not altered their cell tropism (at least in the cell culture system). It would be useful to be able to test the capripoxvirus RPV recombinant vaccines in animals which are normally refractory to infection with capripoxvirus.

2) Lelystad

Transfer plasmids were constructed in which the genes were fused in frame with the gX signal sequence using a plasmid which contains 500 bp of flanking PRV sequences.

Recombinant viruses were generated by homologous recombination. Recombinant PRV was screened for expression of the foreign proteins by using protein specific monoclonal antibodies. Although the F and H genes of RPV were correctly inserted in PRV, expression of neither the F nor the H protein could be detected in cells infected with recombinant virus. However, in cells infected with recombinant PRV that carried the E1 gene, E1 expression could be easily detected. Insertion of E1 into the PRV vaccine strain did not cause any change in cell or host tropism in pigs (the natural host of PRV), rabbits, hamsters, rats, mice and rhesus monkeys. On the contrary, pigs inoculated with the recombinant vaccine strain shed less virus over a shorter period of time than pigs inoculated with the parent vaccine strain, indicating a slight attenuation by the insertion of E1.

Therefore, the virulence and pathogenesis of this 'worst case' recombinant virus (PRV E1+) was studied in pigs and compared with the virulence and pathogenicity of the parent strain. PRV affects the respiratory tract and the central nervous system, whereas HCV has a distinct affinity for cells of the lymphoretical organs. However, no increase in virulence or changes in pathogenesis were detected between the parent strain and the 'worst case' recombinant.
We conclude that foreign genes, when incorporated in PRV, are not always expressed, and that the expression of E1 of HCV did not alter the biological characteristics of PRV.

3) Compton

(a) Recombination between Homologous Poxviruses:

From mixed infections of primary chick embryo fibroblast (CEF) cells, FPV recombinants that were both gpt^+ and $lacZ^+$ were identified in the presence of X-Gal, together with plaque production in the presence of mycophenolic acid. Progeny virus genotypes were confirmed as being recombinants by hybridisation with gpt and lacZ DNA probes. In both the presence and absence of MXH, the observed frequency of recombination was between 3-9%. However, the actual rate of recombination may be up to 100% greater than the observed frequency, due to recombination occurring between the two ITR regions, where the gpt gene is situated, thus resulting in virus with gpt at both or neither ends. Therefore, the frequency of recombination was determined to be 18% for a single two-factor crossover event between homologous poxviruses, using two marker genes separated by approximately 50Kb.
When homologous poxviruses were used to co-infect susceptible cells *in vitro* the progeny virus were represented by both parental genotypes and a mixture of these parental genotypes distinguished by the transfer of marker genes, due to recombination events.The calculated frequency of this recombination, as determined by this set of experiments, agreed with that reported for intergenic recombination between VV mutants.

(b) Recombination between Heterologous Poxviruses:

Both the VV-IBV spike and the FPV -IBV spike-gpt recombinants have been constructed and they have been used in mixed infections of CEFs to see if 'forced recombination' occurs between FPV and vaccinia. Initial results appear to be

positive, in that gpt-positive virus that can replicate on mammalian cells have been isolated. Tests to confirm whether these virus are 'forced' recombinants are underway. Poxvirus-based vectors have been developed as efficient and effective vehicles for delivering foreign genes to elicit protective immune response. At least 25Kb of exogenous DNA can be inserted into vaccinia virus, for instance. Therefore, viruses capable of expressing many different foreign proteins could be engineered for use as polyvalent vaccines, designed to afford protection against different combinations of pathogens found in various geographical areas of the world.

Recombinant poxviruses, represented by vaccinia and fowlpox viruses, have been genetically manipulated to contain relatively small regions of known homology, as might be found in recombinant FPV in which vaccinia virus gene promoter regions are commonly incorporated. Our model recombination experiments indicate the frequency with which recombination can occur between heterologous viruses when only relatively small regions of sequence homology are present. These studies could be extended to determine the smallest region of homology necessary for recombination to occur. The results also cast light on the process of compartmentalisation of replicating virus DNA. Recombination between FPV and VV DNA could not occur if the replicating DNA were held in distinct compartments of the cytoplasm.

At present, our experiments have shown that the frequency of recombination between two homologous viruses could be up to 18%, with that for heterologous viruses being considerably lower. This represents a 'worst case' scenario using *in vitro* models, but would indicate that such events, although extremely uncommon may occur in the field. As random sampling techniques to detect such events would be relatively inefficient, controlled *in vivo* recombination experiments (as detailed in our original proposal), based on the experimental design and reagents used *in vitro*, would be necessary to assess the potential frequency of recombination *in vivo* following mixed infection.

4. Liège

The results of this study have shown that foxes are not very susceptible to cowpox virus which, like recombinant vaccinia, can be isolated from only restricted sites at a low titre and during a short period after experimental inoculation. Thus the possibility of recombination, in foxes, between both viruses should be virtually zero. The very low prevalence of cowpox virus infection in foxes would probably remove this consideration completely..

5. Liverpool

Serological surveys using immunofluoresecence assay (IF) and virus neutralisation assays on sera collected in Great Britain and Belgium identified *Orthopoxvirus* antibody in three main species, namely bank voles (*Clethrionomys glareolus*), field voles (*Microtus agrestis*) and woodmice (*Apodemus sylvaticus*). Antibody was detected in only one house mouse and in none other species tested. It was not possible to isolate virus from any wild animal tissues, thus it is not possible to state that the antibody detected was due to cowpox virus infection. Further studies in laboratory-bred bank voles and woodmice have shown both species to be susceptible to cowpox virus infection. However, bank voles are not susceptible to either ectromelia and vaccinia virus, the other two orthopoxviruses which might be considered alternative causes of the detected antibody.

Orthopoxvirus-free colonies of bank voles, field voles and woodmice were established in the laboratory, although the limited time available meant that further experimental studies were undertaken only with bank voles and woodmice. Bank voles were shown to be susceptible to cowpox virus by footpad inoculation (1 pfu) and by oronasal inoculation (10 pfu). Woodmice were shown to be susceptible to cowpox virus by footpad inoculation (10 pfu). Some evidence indicated that cowpox might cause mild respiratory disease and delayed reproduction in bank voles. Only small amounts of virus were re-isolated from inoculated animals, and generally only from sites of inoculation and for only for a few days after inoculation. This result could explain the negative results obtained with the virus isolation from wild rodents. Tissues are still being processed for histopathology and polymerase chain reaction studies. Further studies with bank voles showed that footpad inoculation with doses of up to 50000pfu vaccinia virus and 5000 pfu ectromelia virus did not cause seroconversion, and no virus was reisolated from inoculated animals. This demonstrates that bank voles are not susceptible to infection with either virus and suggests that the risk of co-infection with vaccinia virus and cowpox virus and the production of recombinant virus is highly unlikely.

Although further work is underway to study the pathogenesis and transmission of orthopoxviruses in bank voles, field voles and woodmice, the data so far collected suggests that

1) The reservoir host(s) of cowpox virus in Western Europe are bank and field voles and woodmice
2) The risk of recombination between cowpox virus and recombinant vaccinia virus in the rodent reservoir is small.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

Pirbright designed, constructed and tested double recombinant virus vaccines which protect cattle, sheep and goats against capripox (lumpy skin disease, sheep/goat pox) and rinderpest and pest des petits ruminants.

Work performed at Liège adds support to the arguments in favour of the use of the vaccinia-rabies recombinant vaccine in large-scale vaccination campaigns in order to eradicate rabies.

Although virus has not yet been isolated from wild rodents in Western Europe, the evidence collected strongly suggests that three small rodent species are the true host of cowpox virus.

MAJOR COOPERATIVE LINKS

Linkage of the group at Pirbright with the Lelystad Laboratory has been extremely useful in the exchange of ideas and techniques covering the molecular manipulation of viruses. In common with all of the participants we have benefited from the annual group meetings and several meetings of one or two of the participating groups.

Linkage and meetings of the group at Leystad with the group at the Pirbright Laboratory has been extremely useful in the exchange of ideas and techniques covering the molecular manipulation of viruses.

Work at Liège was performed mainly in collaboration with the team in Liverpool. The experimental protocol for the first part of this work was established during a meeting in Liverpool. They provided the cowpox virus strain and the protocols. All

serological tests were performed in Liverpool. Sera or other samples collected in Belgium were therefore sent to Liverpool.
Bovine, ovine and wild ruminants sera collected in Belgium and in France were also sent to Pirbright for a capripoxvirus serological survey.

The major collaborative link of the group at Neston was with the group at Liège. However it is interesting to note that projects, at Liverpool, unrelated to EC benefitted from herpesvirus techniques of Lelystad and that the work at Liverpool has led to collaboration with mathematical ecologists at Liverpool and further funding to develop mathematical model of cowpox epidemiology.

PUBLICATIONS

Joint publications

Comparison between the susceptibilities of the red fox (Vulpes vulpes) to vaccinia-rabies recombinant virus cowpox virus. D. Boulanger, B. Brochier, A. Crouch, M Bennett, R.M. Gaskell, D. Baxby, P.-P. Pastoret. Vaccine: (in press)
Serological survey of cowpox virus infection of wild mammals in Belgium. D. Boulanger, A. Crouch, B. Brochier, M. Bennett, R.M. Gaskell, D. Baxby, P.-P. Pastoret. (in preparation).

Individual publications

Recombinant Capripoxvirus expressing the hemagglutinin protein gene of rinderpest virus: Protection of cattle against rinderpest and lumpy skin disease. C.Romero, T.Barrett, R.W.Chamberlin, R.P.Kitching, M.Flemming and D.N.Black. Virology (in press).
Protection of goats against peste des petits ruminants with a recombinant capripoxvirus expressing the fusion and haemagglutinin protein genes of rinderpest. C.Romero, T.Barrett, R.W.Chamberlin, C.Bostock and D.N.Black. Vaccine: (in press)

Identification of genes involved in latency and reactivation of *Pseudorabies* virus, use in biological containment study of viral genomes in pigs (BIOT CT-910297)

COORDINATOR:
A. JESTIN, CNEVA, Ploufragan, F
PARTICIPANTS:
H.J RZIHA, FRCVDA, Tübingen, D
I. ALMEIDA, NVRL, Lisbon, P
P. SHELDRICK, IRSC, Villejuif, F
M. RIVIERE, Rhône-Mérieux, Lyon, F
ASSOCIATED PARTICIPANT TO NVRL:
H. TSIANG, Service Rage I. Pasteur , Paris, F

BACKGROUND INFORMATION

The alphaherpesvirus Pseudorabies virus (PRV, Suid herpes1) has the propensity to establish a reactivable, latent infection in the natural host, swine (not only in neural tissues, but also in lymphocytes). Available vaccines including gene technologically derived life vaccines do not prevent latency, and life vaccines can also establish a latent infection. Recently a limited region of the PRV genome was described which appears to be transcribed predominantly during the latent state of infection. This transcription (LAT = latency-associated transcript) might be regulated by a potential promoter (LAP). The identification and manipulation of LAT as well as LAP will not only help to define its functional role in latency, but might also provide the possibility to construct new safer life vaccines.

Pseudorabies virus (PRV) has a preferential tropism for the nervous system. Primary CNS infection may occur in the context of an acute infection, often lytic or at somewhat later times in the establishment of a latent infection. Despite these observations little is known concerning the mechanism by which the virus initially gains access to the nervous system, which cells are initial targets for PRV infection, and what virus or host factors may facilitate nervous system invasion. Consequently, additional studies are necessary utilizing *in vitro* tissue culture systems and *in vivo* animal models, in order to assess their contribution to the neuropathogenic properties of PRV.

Cytokines are well known to be important mediating factors mediating the host's immune defence against viral infections. Among the cytokines involved, TNFα has been shown to reduce replication and to enhance clearance of the infecting virus in the mouse-vaccinia virus model when its gene is incorporated into the viral genome (Sambhi et al. 1991. *Proc. Natl. Acad. Sci. USA 88, 4025-4029)*, suggesting that such an approach might be usefully applied to herpesviruses, both for studies of pathogenesis and latency in the host and for vaccine development. Pseudorabies virus (PRV), a well characterized neurotropic porcine herpesvirus, was chosen to serve as vector for the porcine TNF-gene.

OBJECTIVES AND PRIMARY APPROACHES

After DNA sequencing and transcriptional analysis, defined parts of LAT were plasmid-cloned and fused to the functional promoterless luciferase gene used as an indicator for cis-acting promoter elements. After plasmid DNA transfection of

different cell lines quantitative measurement of transient luciferase expression demonstrated LAP elements and allowed a delineation of potential promoter(s). A new non essential, non-coding intergenic site was localized on the PRV genome and could be used to insert and express foreign genes. Such a wild-type PRV expressing the functional E. coli lacZ gene, was used for production of LAT-PRV deletion mutants. Comparisons of *in vitro* growth characteristic and LAT activity of the PRV LAT-mutants with wild-type PRV have been performed as well as first *in vivo* experiments in pigs.

Genetic variants of PRV were selected to examine the kinetic parameters of PRV neuropathogenesis. Although PRV infection of continuous human and murine cell lines of neural origin has provided some very important observations regarding cell tropism, cytopathogenicity and replicative properties *in vitro*, this approach does not permit a systematic evaluation of the effect of infection on host cellular events. The use of primary cultures derived from dorsal root ganglia and cerebral cortex of fetuses of 35 days gestational age, provide *in vitro* systems that approximate the interaction of PRV with the swine nervous system.

One goal of this study was to gain more insight into the functional role of LAT in latency. During this phase, the virus is hardly detectable by standard techniques, and quantification of the amount of viruses in a particular tissue was not possible. Therefore we first developped a highly sensitive, specific and quantitative PCR method to detect low amount of PRV virus in infected tissues. Then, this method was used to monitor *in vivo* the amount of virus in animals infected with a PRV LAT-promotor deletion mutant, in comparison with animals infected with a PRV wild-type strain.

The last objective was to develop recombinant PRV carrying the porcine TNFα gene under the control of selected viral promoters. The porcine TNFα-gene was first modified by deleting a putative tissue-specific regulatory sequence from the 3' non-translated region and a TATA-box and several binding sites for other transcription factors from the 5' non-translated region. The resulting recombinant vector contains 55 base pairs upstream of the ATG codon and 300 base pairs downstream of the stop codon.

RESULTS AND DISCUSSION

A. *In vitro* analysis of the LAP

After numerous experiments to establish the optimal transfection protocol for measuring transient luciferase expression in the different cell lines, we found comparable promoter activity in 4 different epithelial cell lines (of porcine, bovine, and simian origin) and mouse neuroblastoma cell lines. A ca. 270 base pair fragment of LAP could be determined to achieve maximum *in vitro* promoter activity. Furthermore, we could show by PRV superinfection experiments that late viral gene products did not transactivate, but rather suppress the activity of PRV LAP. Finally, the results appear to indicate the presence of a second LAP or important enhancing elements some 100 bases upstream of the first LAP.

B. Construction of viral mutants

Starting with the lacZ-labeled wild-type PRV we succeeded in the construction of so far 4 different LAT-deletion mutants. These virus mutants differ in the size of the deleted LAT (and LAP) part. *In vitro* tests demonstrated that these PRV mutants exhibit similar or identical growth characteristics as wild-type PRV. For analysing LAT, a specific RT-PCR was established. First results showed that all

LAT-/LAP-mutants express a LAT RNA spliced identical to wild-type virus. Therefore, these data also indicate the existence of an additional promoter.

C. Analysis in cell culture

Virological studies have been performed, by:

(i) identification of either viral DNA or mRNA by in situ hybridization techniques with cDNA probes labeled by multipriming, or cRNA probes synthesized by reversed transcription;

(ii) identification of viral antigens by immunohistochemistry;

(iii) production of infectious progeny virus.

In every instance the phenotype of infected cell population has been identified by immunohistochemistry with cell type specific markers.

Infection of primary cortical and sensory neuronal cells by PRV resulted in transient expression of viral gB (gII) protein with maximal amounts observed at 3 days post infection. PRV antigenic expression are observed earlier in both fetal cell populations and the antigen appears, distributed diffusely in the cytoplasm. The data indicate that the microglia, Schwann or DRG fibroblastoid cells were productively infected. In conjunction with the detection of PRV-specific protein products, viral mRNA accumulated with a kinetic pattern of expression which parallels that observed for the protein. Neural cell populations infected with wild type PRV strains showed cytopatic changes as early as 18 hours after infection, appearing first in neuronal cell bodies followed by axonal abnormalities.

The release of infectious progeny virus was measured by titration of extracellular and cell-associated virus. Titers in the cultures inoculated with wild type strains increased continuously, there after they declined and were always higher for the cell-associated virus.

D. First animal experiment

A highly specific quantitative PCR method was developed based on the use of an internal standard which is added at known concentrations to the components of the PCR reaction and co-amplified with the target DNA. Both amplified species which differ by a size difference of a few base pairs are then labeled with fluorescent primers and separated by the means of an automated sequencer. It was concluded that 5-10 copies of the PRV genome were quantifiable using this technique. Pigs were infected with a PRV deletion mutant totally lacking the potential LAP (PHY-MS3B) and with wild type virus PHY-B111, respectively, and euthanized various times after infection.

Viral DNA was detected by quantitative PCR for both strains in several neural tissues. From the results obtained so far, the amounts of virus present in latently infected lymphocytes or trigeminal ganglia do not appear to differ significantly between animals infected with either strain.

E. TNF alpha — PRV recombinant

To determine whether the modified gene could express biologically active TNFα it was inserted into a shuttle vector downstream of the RSV LTR promoter and the constuct was intoduced into MDBK cells by transfection. A stable transformed cell was obtained which secreted 20-30 pg/ml of active TNFα into the growth medium after five days in culture. Next, the TNFα gene was placed under the con-

trol of the PRV glycoprotein X promoter and inserted into a vector carrying PRV sequences from the targeted region of the viral genome. In the resulting plasmid, the promoter gX-TNFα gene couple is flanked by several kilobases of viral sequences to facilitate recombinational insertion into the PRV genome.

MAJOR SCIENTIFIC BREAKTHROUGHS

Promoter activity of potential PRV LAP(s) could be shown, and for the first time defined LAT-mutants of PRV have been obtained. Already initiated animal experiments will now enable more insight into the functional role LAT with regard to latency and neurotropism of this alphaherpesvirus in the natural host.

From the methodological standpoint, our major achievement is the development of quantitative PCR method. Using this method, it is now possible to detect and to quantify low amount of virus originating from various latently infected tissues. More recently, some advances have also been made to detect latent PRV virus directly on organ slices by in situ PCR. By using a PRV mutant totally lacking the potential LAT promotor, some informations regarding the possible role of LAT in latency were obtained.

If our preliminary results are confirmed, it appears that the presence of the LAT promotor is not necessary for the establishment of latency and does not influence the amount of virus found in infected animals.

The recombinational insertion plasmid has been successfully used to generate recombinant PRV strains carrying an insertion of the expected size and expressing TNFα mRNA. Experiments to quantify biological activity are currently underway.

MAJOR COOPERATIVE LINKS

Cells, plasmids, recombinant virus, and information concerning experimental results have been freely and frequently exchanged with the Tubingen Laboratory. For *in vivo* experiments the PRV strains were obtained from this lab. Samples from infected animals were sent from Ploufragan.

Several meetings were organized between all parners as well as visits between labs to discuss progress.

PUBLICATIONS

Individual publications

Boutin, P., Arnauld, C., Thiery, R., Costa, J.M., Vidaud, M. and Jestin, A. Chemiluminescent detection of amplified pseudorabies virus gop50 DNA with immobilized probes on microtitens wells. *Acta Veterinaria Hungarica.* 1994, **42**, 387-385.

Thiery, R., Boutin, P., Arnauld, C. and Jestin A. Pseudorabies virus latency: a quantittive approach by polymerase chain reaction. *Acta Veterinaria Hungarica.* 1994, **42**, 277-287.

Foulon, T., Kaidot, T., Fuchs, W., Rziha, H.J. and Sheldrick, P. A bovine cell line stably expressing porcine tumor necrosis factor alpha (TNF-α): growth properties and permissivity for pseudorabies virus replication. *Acta Veterinaria Hungarica.* 1994, **42**, 195-203.

Biosafety of genetically modified baculoviruses for insect control (BIOT CT-910291)

COORDINATOR:
J.M. VLAK, Agricultural University, Wageningen, NL
PARTICIPANTS:
R.D. POSSEE, NERC Institute of Virology, Oxford, GB
J. HUBER, Institute for Biological Control, Darmstadt, D

BACKGROUND INFORMATION

Baculoviruses are insect pathogens which are successfully used as biological control agents of insect pests in agriculture and forestry. Safety testing has confirmed that these viruses are insect specific and cause no hazard to other animals, beneficial insects, or to plants. Despite their use for over five decades, no adverse side-effects have been reported. However, knowledge on the baculovirus ecology is scarce and could be elucidated by using genetically marked viruses. These have been provided through the EC-BAP programme.

More widespread use of baculoviruses has been hampered by their relatively slow speed of action, limited host range and high cost of production. The low speed of action, in particular, which is most relevant for crops with low damage thresholds, called for improved baculovirus strains with enhanced insecticidal activity. A number of potentially useful baculoviruses (*Autographa californica* nuclear polyhedrosis viruses or AcNPV) with enhanced insecticidal activity have been obtained, including those containing insect specific toxin genes or having a deletion in the ecdysteroid UDP glucosyl transferase (egt) gene. Using baculoviruses with these traits, the time giving 50% mortality of insect larvae is reduced by one or two days in a process which otherwise takes between 5-10 days.

There is limited knowledge on the biosafety of genetically modified baculoviruses and it is unlikely that all questions raised can be satisfactorily answered without understanding the behavior of baculoviruses in the environment. Information on the parameters which influence baculovirus infections in the field allows predictions to be made about the behavior of genetically modified baculoviruses in the environment. Risk assessment involves the identification and analysis of possible risk factors. The management of potential risks (defined as exposure x hazard) can be based on a reduction of exposure. i.e. of dispersal, persistence or yield of genetically modified baculoviruses.

OBJECTIVES AND PRIMARY APPROACHES

The primary objectives of the project have been to generate and evaluate genetically modified baculoviruses (AcNPV) with reduced capacity for survival and develop techniques in order to predict their fate and interaction with non-targets in the field. The primary approach have been the construction of AcNPV mutants either with deletions of genes thought to be involved in persistence, spread and survival or mutants with a 'built-in' suicide mechanism. The availability of such mutant baculoviruses allows a comparison of wild-type (*wt*) and genetically modified baculoviruses in terms of their competitiveness and behavior in permissive, semi-permissive and non-permissive insects.

RESULTS AND DISCUSSION

A. Generation of deletion and 'suicide' mutants

AcNPV contains at least six genes that, in principle, could affect their persistence, survival and spread in the environment. The polyhedrin gene (*ph*) codes for the matrix protein of polyhedra, which are important for convenient application of virus in the field. The *egt* gene is responsible for control of the larval molt; deletion of this gene results in premature molt, increased speed of kill and reduced virus yield. The pp34 gene (*pe*) is involved in the formation of the polyhedral envelope. Its deletion results in envelope-less polyhedra, which have an increased sensitivity for weak alkali and which are likely to have reduced persistence in the field. The p10 gene (*p10*) is responsible for the formation of fibrillar structures in infected cells and for the disintegration of nuclei, resulting in impaired release of polyhedra in the field. The chitinase gene (*chit*) cause the rupture of insect cuticle, whereas the product of the cathepsin gene (*cath*) proteolytically cleaves the internal proteins, both enhancing the release of progeny polyhedra. Deletion of these genes does not effect virus replication or efficacy in insects. Thus, deletion mutants in the *p10, chit* and/or *cath* gene have a reduced propensity for spread in the environment.

Single, double and triple deletion mutants involving the *egt, p10, pe, chit* and *cath* were engineered, with and without beta-galactosidase (LacZ) as reporter gene. Recombinants were engineered in a two-step process: first gene interference with LacZ gene (blue) and then replacing LacZ with a deletion (white). The promoters of the deleted genes were kept intact in order not to disturb overall transcription during virus replication. The replication of the deletion mutants *in vitro* was unaffected. The majority of these mutants were tested for biological activity. One of the deletion mutants (AcNPV/p10-LacZ+) were analyzed in a microcosm.

Attempts were made to engineer a 'suicide recombinant by placing the *Escherichia coli* lac operator (*lacO*) upstream or downstream of the polyhedrin transcriptional start site and the lac repressor (*lacI*) under the control of the *Drosophila* heat shock promoter hsp70. In principle the expression of polyhedrin and hence the formation of polyhedra is blocked in the field. In the presence of IPTG for (example under production conditions) the block will be released, since the *lac* repressor preferably binds to IPTG. Using a transient reporter systems it appeared that the introduction of the *lac* operator sequence (17-mer) up- or downstream of the polyhedrin promoter was inadequate to block the expression of the polyhedrin gene. Further experiments are required to exploit the potential of this approach.

B. Competitiveness and fitness of wild-type and recombinant baculoviruses in insects

UV-persistence was studied by inactivation of dry virus deposits by direct exposure to artificial UV-sunlight. Comparison between the *wt* AcNPV and an AcNPV/*pe*-minus mutant showed no significant difference in half life. The loss of the polyhedron envelope apparently does not result in a increased UV sensitivity.

Comparison of six AcNPV deletion mutants (without the LacZ gene) in a series of bioassays in 2nd instar cabbage looper (*Trichoplusia ni*) larvae (a sensitive host) showed that the single deletion mutants did not differ significantly from wt AcNPV in terms of infectivity (LD_{50}). There were indications that a double deletion mutant (AcNPV/*egt*-minus/*p10*-minus) might be less infective.

One of the key questions in terms of whether a genetically modified virus will survive and spread after release into the environment is whether this modification has altered its fitness and how it interacts with other wt viruses it encounters. A technique was developed, using AcNPV-LacZ, whereby mixed infections could be followed through multiple passages in insect larvae and parameters such as replication rate and yield could be measured with precision. Preliminary single infection assays had already indicated that Ac/LacZ and a construct which expressed an insect-selective scorpion toxin (AcNPV/AaIT) produced less polyhedra than wt AcNPV. Mixed infections of wt AcNPV and AcNPV/LacZ in different proportions (1:10, 1:1, 10:1) were passaged in cohorts of insects. After each passage the proportion of each virus type in the mixture was estimated and the remainder of the virus re-passaged. After several passages it was found that AcNPV/LacZ was rapidly lost from the system, suggesting reduced biological fitness of the recombinant.

These assays gave some indication of what might happen if two viruses were introduced synchronously. However, a situation may also arise in the field where one virus is ingested after another. Experiments similar to those described above were set up using *wt* AcNPV and AcNPV/LacZ, but where one virus was introduced a set of time after another (48 h for instance). The results showed that the first virus to infect does not prevent the infection by a second virus, but does interfere with its replication, significantly reducing its titre. These results have important implications for risk assessment as this interference effect means that the opportunities for recombination between genetically modified baculoviruses and other *wt* baculoviruses will be considerably reduced where infection with another virus is asynchronous.

C. Host range and biological activity of genetically modified baculoviruses in permissive, semi-permissive and non-permissive hosts.

The biological activity of *wt* AcNPV a AcNPV/*p10*-minus recombinant expressed as LC_{50}-value (lethal concentration 50%) was compared in standard bioassays. Five concentrations of virus were used against first instar larvae of *Spodoptera exigua* and *A. gamma*. No statistically significant differences in infectivity was determined between *wt* and recombinant AcNPV.

Interaction with non-target Lepidoptera is one of the main areas for concern in relation to the biosafety of genetically improved baculovirus insecticides. The AcNPV-LacZ marker system was used to examine the extent of infection in selective permissive (*T. ni*), semi-permissive (*Mamestra brassicae*) and poorly permissive (*S. littoralis*) hosts. This system provided an elegant method whereby the blue coloration generated by the expression of LacZ gene could be monitored in insect larvae itself using both plaque assay and microscopic techniques to study the time course of infection at different doses. In the semi-permissive host, dose/instar combination was shown to be the key in determining the level of infection, mortality only occurring at certain dose threshold. This can be interpreted as being a result of the degree of containment provided by the insect gut: as the dose increases (or the size of caterpillar decreases) the gut is less able to fight off baculovirus infection by sloughing off cells and so the infection is able to gain entry to the rest of the body. One interesting an novel finding was that the trachea appear to be the main conduit for the dissemination of infection through the insect body. At the dose/instar combinations used, no mortality was obtained in *S. littoralis* larvae. However, replicating virus was found in the tracheal cells through to the pupal stage.

D. Release of genetically modified AcNPV in a microcosm

Prior to field release genetically modified AcNPV carrying deletions were tested in a contained ecosystem (microcosm) to obtain data about their behavior, in particular with respect to yield and dispersal. Two deletion mutants were extensively tested in the course of this project: AcNPV/*pe*-minus and AcNPV/*p10*-minus, carrying a deletion in the polyhedral envelope and p10 gene respectively. Both mutants carried a LacZ gene. This reporter allowed the detection of mutant virus in infected larvae.

The microcosm, constructed in the framework of BAP, was separated by an insect screen into two compartments to prevent larval migration from one chamber to the other, but allowing air circulation. Each compartment contained soil as substrate, sugar beet plants and second instar larvae of *S. exigua.*

Wt or recombinant AcNPVs were applied in a 0.4 m^2 compartment at a single dose of 4 x 10^7 polyhedral inclusion bodies (PIBs), equivalent to 10^{12} PIBs per hectare. After one larval generation (25 days), soil, plants and infiltration water were collected from the ecosystem and were bio-assayed for virus quantitation.

Equal quantities of the *wt* AcNPV and AcNPV/p10-minus) applied to 200 second instar larvae of *S. exigua* showed no difference in response to infection (82 and 81% mortality). Budgetary studies of NPV production and partitioning, involving foliage and soil extractions indicated that in the case of AcNPV/*p10*-minus the yield and the distribution over the various sectors (plants 5%, soil 95%, drainage 0.03%) were the same, but that the *wt* AcNPV was 5 times more abundant in the non treated plots, suggesting limited spread of AcNPV/*p10*-minus.

Equal quantities of both viruses applied to 100 larvae gave only minor differences in response (39 and 48% larval mortality, respectively). Small differences in virus distribution between the components of the microcosm were detected between *wt* AcNPV and AcNPV/*pe*-minus. Both viruses were washed into the soil but only the *wt* AcNPV penetrated the whole soil profile of 15 cm depth and led to a contamination of the infiltration water, whereas the recombinant virus was only detected to a depth of 0-6 cm. This suggest a higher retention of recombinants in the upper layers of the soil.

MAJOR SCIENTIFIC BREAKTHROUGHS

1. Generation of stable genetically modified baculoviruses with multiple deletions while maintaining their major biological properties (specificity, biological activity).
2. Successful testing of the behavior of single deletion mutants in a microcosm showing reduced spread.
3. Demonstration that genetically modified baculoviruses may be less competitive the wild type viruses by intereference in asynchronous (common in nature) mixed infections
4 Identification of the role of trachea in the spread of virus infection in the insect body

MAJOR COOPERATIVE LINKS

Free exchange of materials and experimental expertise took place between the participating laboratories. For example, the deletion mutants generated in the Wageningen laboratory were tested for their biological activity in Oxford and

tested in a microcosm in Darmstadt. Regular meetings between participants occurred either planned during scientific conferences or *ad hoc.*

PUBLICATIONS

Joint publications

Bonning, B.C., Roelvink, P.W., Vlak, J.M., Possee, R.D. and B.D. Hammock. 1994. Superior expression of juvenile hormone esterase and beta-galactosidase from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus compared to the p10 protein and polyhedrin promoters. Journal of General Virology **75**: 1551-1556.

Gloger, U., Undorf-Spahn, K., Huber, J. and J.M. Vlak. 1992. UV-persistence of a wild-type and a recombinant *Autographa californica* nuclear polyhedrosis virus. Abstract Book of the 1st EC Sectoral Meeting on Biosafety, Wageningen, The Netherlands, p. 82.

Mans, R.W.M., Hauxwell, I.C., Undorf, K., Possee, R.D., Huber, J., Cory, J.S. and J.M. Vlak. 1993. Engineering for enhanced biosafety of recombinant baculoviruses. Abstract Book of the II International Symposium on Molecular Insect Science, p. 103.

Roelvink, P.W., Van Meer, M.M.M., De Kort, C.A.D., Possee, R.D., Hammock, B.D. and J.M. Vlak. 1992. Dissimilar expression of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus polyhedrin and p10 genes. Journal of General Virology **73**: 1481-1489.

Undorf-Spahn, K., Huber, J. and J.M. Vlak. 1992. Biological behaviour of a genetically modified *Autographa californica* nuclear polyhedrosis virus. Abstract Book of the 25th Annual Meeting of the Society for Invertebrate Pathology, Heidelberg, Germany, p. 141.

Undorf-Spahn, K., Huber, J. and Vlak, J.M. 1992. Simulated field release of a recombinant *Autographa californica* nuclear polyhedrosis virus in a model ecosystem. Abstract Book of the 1st EC Sectoral Meeting on Biosafety, Wageningen, The Netherlands, p. 82.

Undorf-Spahn, K., Huber, J. and Vlak, J.M. 1992. Competitive multiplication and dispersal of a recombinant and wild-type *Autographa californica* nuclear polyhedrosis virus in a model ecosystem. Abstract Book of the Final Sectoral Meeting on Biosafety, Granada, Spain, p. 109.

Vlak, J.M., R.W.M. Mans, J. Huber, K. Undorf, U. Glober, R.D. Possee, I.C. Hauxwell and J.S. Cory. 1993. Biosafety of genetically engineered baculoviruses for insect control. Progress Report BRIDGE (A. Vassarotti and U. Kirchheim, eds.), p. 413-416.

Vlak, J.M. 1992. Engineering of biosafe baculovirus recombinants for insect control. Abstract Book of the XIX International Congress of Entomology, Beijing, China, p. 135

Individual publications

Hauxwell, I.C. 1994. Time course of infection and comparative pathology of *Autographa californica* nuclear polyhedrosis virus in permissive, semi-permissive and poorly permissive hosts. Abstract Book of the International Colloquium on Invertebrate Pathology and Biological Control, Montpellier, France, p. 170.

Hauxwell, I.C., Bishop, D.H.L. and J.S. Cory. 1994. Competition and interference during mixed infections of wild type and genetically modified *Autographa californica* nuclear polyhedrosis virus in an insect host. Abstract Book of the International Colloquium on Invertebrate Pathology and Biological Control, Montpellier, France, p. 161.

Hauxwell, I.C. and J.S. Cory. 1993. Ecologically-based risk assessment of genetically modified baculoviruses. Abstract Book of the Final Sectorial Meeting on Biosafety, Granada, Spain, p. 111.

Hawtin, R.E., King, L.A. and R.D. Possee. 1992. Prospects for the development of a genetically engineered baculovirus insecticide. Pesticide Science **34**: 9-15.

Hauxwell, I.C., Possee, R.D. and J.S. Cory. 1992. Genetically engineered baculoviruses and gene survival. Abstract Book of the 1st EC Sectoral Meeting on Biosafety, Wageningen, The Netherlands, p. 81.

Huber, J. and Undorf-Spahn, K. 1992. Simulierte Freisetzung von gentechnisch veränderten Kernpolyederviren von *Autographa californica* in einem Modell-Ökosystem. Jahresbericht Biologische Bundesanstalt für Land- und Forstwirtschaft, p. 101-102.

Huber, J., Undorf-Spahn, K. and Gloger, U. 1992. Vergleich der UV-Empfindlichkeit eines Kernpolyedervirus von *Autographa californica* in seiner Wildform und als genetische Rekombinante. Jahresbericht Biologische Bundesanstalt für Land- und Forstwirtschaft, p. 102.

Mans, R.M.W. and Vlak, J.M. 1992. Engineering of biosafe recombinant baculoviruses. Abstract Book of the 25th Annual Meeting of the Society for Invertebrate Pathology, Heidelberg, Germany, p. 139.

Mans, R.M.W. and Vlak, J.M. 1993. Engineering of biosafe recombinant baculoviruses. Abstract Book of the Final Sectorial Meeting on Biosafety, Granada, Spain, p. 108.

Mans, R.M.W. and Vlak, J.M. 1992. Engineering of biosafe recombinant baculoviruses. Abstract Book of the 1st EC Sectoral Meeting on Biosafety, Wageningen, The Netherlands, p. 80.
Possee, R.D., Whitehouse, A., Pullen, J. and Lopez-Ferber, M. 1993. The development of biosafe baculovirus insecticides with inducible polyhedra production. Abstract Book of the Final Sectorial Meeting on Biosafety, Granada, Spain, p. 110.
Vlak, J.M. 1993. Genetic engineering of baculoviruses. In: Opportunities for Molecular Biology in Crop Production (D.J. Beadle, D.H.L. Bishop, L.G. Copping, G.K. Dixon and D.W. Hollomon, eds.), British Crop Protection Council, p.11-22.
Vlak, J.M. 1994. Novel strategies in the development of engineered baculovirus insecticides and concepts for biosafety. Proceedings of the VIth International Colloquium on Invertebrate Pathology and Microbial Control, Montpellier, France, p. 469-501.

Risk evaluation for genetically modified microbial inoculants (BIOT CT-910283)

COORDINATOR:
M.P. NUTI, University of Padova, Padova, I
PARTICIPANTS:
M. BASAGLIA, Heligenetics, Gaiba, I
J.E. BERINGER, University of Bristol, Bristol, GB
M.F. O'GARA, University College, Cork, IRL

BACKGROUND INFORMATION

Irrespective of the method used for strain selection and improvement (rDNA or conventional genetic), there is a lack of knowledge on how a given microbial inoculant will survive, persist, move away from inoculum site or will be affected by environmental and agronomic factors. Monitoring the fate of a microbe, released in an open environment, might be difficult particularly when low density populations become established in soil. This will be of crucial importance if a 'risk assessment' or an environmental impact analysis is required for regulatory reasons, such as those imposed by the European Directive EC/220/90.

OBJECTIVES AND PRIMARY APPROACHES

The project is targeted at better understanding how (genetically modified) microorganisms to be used as soil microbial inoculants can persist, multiply and spread in their environment. The different tasks of the project include
(a) the development of tools for identification and characterization of rhizobial inoculants,
(b) the identification of factors influencing the survival of rhizobial inoculants,
(c) the study of the fate of genetically modified rhizobia during inoculant manufacturing and storage, as well as in the field, and
(d) development and evaluation of intrinsically biosafe soil microbial inoculants.

RESULTS AND DISCUSSION

1. Padova

For the construction of reporter systems, a catabolic gene (coding for β-galactosidase) and a positively selectable gene (coding for resistance to mercury chloride) were chosen as 'reporter' traits. Mercury resistance can be particularly useful when dense soil suspensions (i.e. 10^{-1}-10^{-2} dilutions) have to be plated for cfu (colony forming units) determination; fungal growth is minimized by incorporating the chemical in the medium without affecting the growth of the rhizobial strain. The *lac*Z gene comes from *Escherichia coli*; its promoter was removed and replaced with a synthetic fragment which proved to possess a strong activity as promoter in *E. coli* and in *Rhizobium leguminosarum*. A *lac* operator was added between the promoter and the *lac*Z gene to allow inducible expression; the operator can also be removed, leading to constitutive gene expression. A *lac*I^q gene from *E. coli* was found necessary, as *R. leguminosarum* does not show any *lac*I-like function able to act on the operator. All these structures are assembled into a cartridge, about 10 kb, containing unique restriction sites at the extremities. The lac-mer cartridge was inserted into the broad host range vector pRL497, a RSF1010 derivative (incQ). The resulting plasmid pDG3 is about 16.5 kb and is

devoid of any antibiotic resistance. Plasmid pDG4 is identical to pDG3, except for the absence of the lac operator. This results in the constitutive, high level expression of *lac*Z. Plasmid pIRL1 was constructed by inserting the lac-mer cartridge into pMS102, a pSUP102 derivative containing a fragment of *R. leguminosarum* DNA which carries the *rec*A gene sequence. The cartridge is inserted within the coding region thus inactivating *rec*A. Since pIRL1 is not able to replicate in *Rhizobium*, the reporter gene can be maintained in the host only if homologous recombination occurs. In this way it is possible to have a copy of the reporter cassette located on the chromosome along with the inactivation of the rec functions.

A method was also developed for the study of resident *Rhizobium* populations. The method, based on the analysis of DNA profiles by pulsed-field gel electrophoresis in contour-clamped homogeneous electric field (PFGE-CHEF), allows to discriminate strains having similar plasmid profiles.

Using genetically modified *R. leguminosarum* containing the various reporter gene cassettes described above, a study was carried out, aimed at verifying the fate of GM rhizobia during inoculant manufacturing and storage, along with the stability of the genetic modification. Three inoculant formulations were chosen for the risk assessment study, i.e. liquid (Rizoking™), adsorbed onto sterile vermiculite (Goldcoat™) and adsorbed onto non-sterile peat.

Unexpected behaviour of GM strains with respect to the unmodified parental strain was not observed. This is irrespective of the localisation of the genetic modification, i.e. chromosomal or plasmid-borne. However, different population dynamics could be described, depending upon the intrinsic stability of the construct, i.e. regulated or unregulated highly unstable cartridge.

2. Gaiba

Two sites used in the BAP programme, have been identified. Cropping history and MPN counts were recorded and sent to Bristol. Nodules from the plants were collected and sent to Bristol for analyses.

The general objective was to evaluate the behaviour of genetically modified Rhizobium strains used for inoculant production, during preparation and storing phases.

The production of the inoculants was carried out in the lab of Heligenetics, using four strains previously described (Giacomini A., Ollero F.J., Squartini A. and Nuti M.P. 1994 Construction of multipurpose gene cartridges based on a novel synthetic promoter for high level gene expression in gram-negative bacteria. Gene, in press), in particular:

Rhizobium leguminosarum bv. viciae 1003 (wild type)
Rhizobium leguminosarum bv. viciae 1110 (1003 containing the plasmid pDG3)
Rhizobium leguminosarum bv. viciae 1111 (1003 containing the plasmid pDG4)
Rhizobium leguminosarum bv. viciae 1112 (1003::pIRL1, containing pIRL1 integrated in the chromosome).

The strains, supplied by Prof. Nuti in Padova, were cultured in small scale fermentors (10 l) and rhizobial inoculants were produced. Three inoculant formulations were choosen for the risk assessment study, i.e. liquid (Rizoking™, absorbed onto sterile vermiculite (Goldcoat™ and absorbed onto non-sterile peat.

After production, the samples were tranferred to prof. Nuti's lab in Padova where the monitoring of the strains was carried out.

When studying the survival of Bradyrhizobia into soil in northern Italy, no simple correlation with the range of standard physico-chemical characteristics of the soil could be found. Among possible factors influencing the persistance of a given strain, in the River Po Valley in northern Italy, the role of water table variation could be ascertained. It was of interest to evaluate the population dynamics of *Bradyrhizobium japonicum* 902A released as soil liquid inoculant in the presence/absence of its host plant and low/high water table level. A field (a calcareous loamy soil) was selected in the eastern pianura Padana. The resident *B. japonicum* population was less than 0.6 cell/g soil dry weigh. Two plots of 80 m^2 were prepared. In one plot, the water table level was weekly adjusted (for 23 weeks) by a drip irrigation system raising it from 180 cm to about 90 cm. Subplots were sown with inoculated and uninoculated soybean (40 seeds per m^2). MPN counts were performed 1, 3, 7, 10 and 12 months after sowing. In subplots where host plants was present and the water table was artificially raised, 12 months after sowing, the residual population was 85% less compared to subplots where water table was not adjusted; in subplots inoculated only with rhizobia, the residual population was 34% less compared to subplots where the water table was mantained at 180 cm depth.

The result strongly indicate that water table adjustments affect the survival of the introduced population of *B. japonicum* during the first 12 months of the experiments both in the presence and in the absence of the host plant.

3. Bristol

PCR fingerprinting allowed us to demonstrate that isolates were derived from inoculant strains. In one case we were able to demonstrate that information about an inoculant was incorrect because the strain isolated from the field showed no relationship to the expected strain. PYMS provided a useful second check on strain identity. Becaus PYMS is an assessment of similarity on the basis of the phenotype of the organism, we were able to show that in one site the isolates represented two populations derived from the original strain. We assume that this represents genetic selection for organisms better adapted to some niches in that soil. It is also possible, but unlikely, that a mixed population was added at the time of inoculation. Our work with *B. japonicum* has shown us that it is relatively simple to characterise soil populations once appropriate primers and conditions are available for the PCR reaction. PYMS, wich is expensive, provided a useful backup for a more detailed analysis of chosen isolates. PYMS can detect change which might have arisen from the expression of only one gene if this affects the composition of molecules in the organism. We think that the most likely changes in the soil will be those arising from changes in polysaccharide production, and are therefore not surprised that PYMS showed a change in the population at one site.

4. Cork

a. Development of alternate vector systems

(i) The natural selection system for *thy* mutants is based on the ability of such mutants to grow in the presence of antifolate drugs and an exogenous supply of thymidine. This procedure could be readily employed to select *R. meliloti thy* mutant strains. This method however, failed to provide a selection for similar mutants from *Bradyrhizobium* and *Pseudomonas* spp. Potential *thy* mutants were obtained from a number of *R. meliloti* strains and other *Rhizobium* spp. No additional work has been done with these latter strains.

Unlike the *thy* mutant strains in *E. coli*, we observed that *R. meliloti thy* mutants (once identified) could not be propagated on the selective medium, containing the antifolate drug. *R. meliloti thy* mutants can be readily isolated as single colonies from a lawn of parent cells, but cannot be propagated on fresh selective medium. We have been unable to identify the additional requirement, besides thymidine, for the propagation of these cells in the presence of antifolate drugs. This is a significant disadvantage in working with *R. meliloti thy* mutant strains, since it is not possible to ensure strain purity by culturing in the presence of antifolate drugs and thymidine.

(ii) The thy vector system consists of a free replicating, broad host range plasmid (pG DT10, Tc^R or pGDT11), containing the *Lactococcus lactis thy* gene and a *R. meliloti* host strain deficient in thymidylate synthase. This results in a vector-host combination which is intrinsically stable. Both plasmids are identical, except for the tetracycline resistance gene which is deleted in pGDT11. Therefore pGDT11 is devoid of antibiotic resistance genes. Both plasmids contain a constitutively expressed *lacZY* 'marker' gene. In laboratory scale, hydroponic experiments with pGDT10, plasmid loss from the wild-type host during symbiosis amounted to over 90% over a 5 week period. Plasmid stability in the *thy* mutant background remained virtually 100% (Dowling et al, 1992). These experiments demonstrated that the thy system could dramatically improve plasmid stability in systems where external selective pressure cannot be sustained. The effect of the thy vector system on the competitiveness of *R. meliloti* inoculants was investigated in a mixed inoculum experiment on hydroponically grown Alfalfa plants. Alfalfa seedlings were inoculated with an equal mixture of the RM42 parent strain (lac-) and the RM42T/pGDT11 thy vector system (lac+). The inoculant contained 48% *lac-* versus 52% *lac+ bacteria.* After 4 weeks, this ratio shifted slightly towards the wild-type (i.e. 64% lac- versus 36% lac+ bacteria) A similar ratio was found on the root surface, suggesting that there is free movement of bacteria between the root surface and the rooting solution. Bacteria were reisolated from the nodules. Nodules from only two plants were screened. Plant A showed a ratio of 82% lac- versus 12% lac+ bacteria. Plant B was 100% nodulated by the wild-type strain. This dramatic shift towards the wild-type in nodule occupation suggests that the *thy* vector system impaires the competitiveness of *R. meliloti* strains for nodulation of Alfalfa plants.

(iii) During the summer of 1991, the stability of the thy vector system for *Rhizobium meliloti* inoculants, in conjunction with the lucerne cultivar 'Europe', was evaluated under field conditions. Permission for the environmental release of a geneticaly modified *Rhizobium* inoculant strain had been obtained from the National Recombinant DNA Committee. Plasmid pGDT11 was chosen for this release, because this plasmid vector is devoid of antibiotic resistance genes and contains a constitutively expressed *lacZY* 'marker' gene for monitoring purposes. Sowing took place on the 26th of July 1991. Six square meter plots were used. Plot A was the control and remained uninoculated. Plot B received an inoculation with the parent strain RM42. Plot C received RM42, containing the vector plasmid pGDT11. Plot D received an unrelated inoculant GMI42. Plot E was inoculated with the *thy* mutant strain RM42T. Finally, plot F was inoculated with the vector system RM42T/pG DT11. The size of the inoculums was between 10^{10} and 10^{11} bacteria/m^2. No adverse weather conditions or plant diseases were encountered during the course of the experiment. At the end of the growth season, plants were removed and transferred to the laboratory. Bacteria were reisolated from the nodules. As was evident from their lac phenotype, plasmid stability in a *thy*-host (Plot F) was found to be 91%. This compares with a plasmid stability of 67% in

the parent strain (Plot C). Plant yield was not adversely affected by the presence of the vector system. This demonstrates the effectiveness of the *thy* vector system in ensuring plasmid stability under field conditions.

b. Construction of suicide system for *R. meloti*

We have been concentrating on the development of a containment system for *Rhizobium* spp., based on the limited survival characteristics of a *Rhizobium meliloti* strain mutated in thymidylate synthase and conditional expression of the *thy* gene. Two aspects were considered essential for the success of a containment system. First, the *R. meliloti thy* mutant must be stable and unable to revert back to wild-type. A reverse genetics approach has been used to generate insertion and deletion mutants in *R. meliloti* strains and possibly other *Rhizobium* species. Second, we have to assess the capability of the *thy* mutants to survive in the soil, in the presence and absence of the host plant. The latter was also important to decide on the nature of the promoter that will be used to conditionally express the *thy* gene.

(i) The survival of the wild-type and mutant strains in rooting solution was monitored. After 35 days, significant numbers of the wild-type strain were found to survive in rooting solution, either in the presence or absence of thymidine. In contrast, *thy* mutant strains only survived when thymidine was present in the rooting solution. Also, in the presence of thymidine, the mutant was found to be 100% stable. In the presence of the host plant, the *R. meliloti thy* mutant strain was found to survive the entire length of the experiment (35 days) on the root surface, without the addition of thymidine to the rooting solution. This indicated that the alfalfa host plant can support the *thy* mutants on their root surface. This may indicate the presence of thymidine in the root exudate. The relatively high revertion rate of 26% reverted back to wild-type, indicated that the amount of thymidine supplied is probably only marginal. In contrast to the alfalfa seedlings, sugarbeet seedlings were shown to be unable to sustain the *thy* mutants on their root system. The nodulation of the alfalfa plants by the *thy* mutants, in absence of thymidine was severely delayed. Nodule occupancy showed that 100"% of the *R. meliloti* bacteroids are reverted to wild-type. This indicates that the *R. meliloti thy* mutant strain cannot nodulate, even though it can survive on the root surface of the alfalfa plants. In contrast, when thymidine is provided in the rooting solution nodulation proceeds and no revertion is evident among the reisolated bacteria. We have also consistently observed a reduction in the nodulation rate of the wild-type strain when thymidine is provided in the rooting solution.

(ii) The survival characteristics of the *thy* mutant strain indicated that a promoter which is expressed during symbiosis (and maybe also during nodulation) may be required to conditionaly express the *thy* gene and render the contained *R. meliloti* strains viable and effective. Promoters of genes involved in symbiotic N_2 fixation (*nif*), nodulation (*nod*) and C_4-dicarboxylate transport (*dct*) are being evaluated for this purpose. The *nifH* and especially the *dctA* promoter have been studied extensively in our laboratory. The *Lac. lactis thy* gene can be conveniently cloned as a *Sau*3A fragment into the *Bam*Hl site of pGD926 derived gene fusions. The reading frame at the *Sau*3A site at the third amino acid of the *Lac. lactis* thy gene and the *Bam*Hl site at the beginning of the *lacZY* genes in pGD926 are such that this cloning results in an in-frame gene fusion of the upstream symbiotic promoter to the *thy* gene. pM8210 (*nifH*) and pCU700 (*dctA*, Wang et al, 1993), are both based on the pGD926 plasmid and are being used to fuse these genes to the *L. lactis thy* gene.

(iii) The *R. meliloti thy* gene has been cloned by complementation of a spontaneous *R. meliloti thy* mutant. The gene has been subcloned and mapped to a 1.5 Kb fragment. The smallest subclone, pRMT2, was found to be unstable and spontaneously deleted to form a plasmid pRMT2δ. This plasmid had lost one half of the pUC multiple cloning site, leaving a single *Eco*RI restriction site in the *thy* gene suitable for insertion mutagenesis. Three different mutations have been constructed:
Mutant 1 has a gene cassette coding for a kanamycin resistance gene, cloned into the *Eco*RI site, thus disrupting the *thy* gene.
Mutant 2 is a frame-shift mutation, created by site directed mutagenesis of the *Eco*RI restriction site.
Mutant 3 was obtained by deletion of an *Eco*RI-StuI fragment from the gene.
These mutated *thy* genes will be marker exchanged into a range of *R. meliloti* strains and possibly in a number of related *Rhizobium* species. To date, attempts to clone the *B. japonicum thy* gene through complementation of an *E. coli thy* mutant have been unsuccessfull.

c. Evaluation of a genetically modified inoculum strain

Plasmid pCU2B contains a copy of the gene, which may increase the metabolism of *Rhizobium* bacteroids during symbiosis. This plasmid was stabilised through the introduction of a copy of the *thy* gene (pCU2BT) and introduced in a *thy-* host strain (CM21). Preliminary laboratory scale experiments with CM21/pCU2BT indicated a significant increase in plant yield, when alfalfa plants were nodulated with this GM strain. it was decided to evaluate the performance of this genetically modified organism in soil, under greenhouse conditions. The experiment was set up and the plants were cut every 4 weeks. No significant differences were observed between the various treatments, including the control experiment. Consecutive cuttings, designed to deplete the soil of nutrients, in particular nitrogen content, did not result in a significant difference in yield from the various treatments. In the final analysis it was found that the soil chosen for this experiment (from a field in Youghal Co. Cork) contained a resident *Rhizobium* strain able to outcompete the inoculant strains. Time constraints prevented us from repeating this experiment with Fota soil, which has been shown to be free from resident *Rhizobia* capable of nodulating lucerne plants.

MAJOR SCIENTIFIC BREAKTHROUGHS

The *thy* vector system was developed at Cork and tested under greenhouse conditions. The system was demonstrated to dramatically improve plasmid stability during symbiosis where external selective pressure cannot be sustained. In competition experiments with the wild-type parent strain, the *thy* vector system was found to decrease nodulation efficiency of the host strain. Symbiotic efficiency was indistinguishable from that of the parent strain.

The effectiveness of the *thy* vector system was tested under field conditions. Permission was obtained from the Recombinant DNA commission for an environmental release of the genetically modified *Rhizobium meliloti* inoculant strain. This was the first environmental release of a GMO in Ireland. Under field conditions, plasmid stability was improved significantly by the *thy* system. Plant yield was not adversely affected by the presence of the vector system. This demonstrates that the *thy* vector system is suitable for introducing genes of interest into *Rhizobium* inoculant strains and ensure stable maintenance of the genetic trait under field conditions.

The survival characteristics *R. meliloti thy* mutants indicate that conditional expression of the *thy* gene is a possible method for containment of *Rhizobium* inoculant strains. *R. meliloti thy* mutants were shown to be able to colonise, but not to nodulate the alfalfa host plant. This indicates that a promoter which is expressed during symbiosis (such as many *nif* or *fix* promoters), may be suitable for the construction of a containment system for *Rhizobium* inoculants. Conditional expression of the *thy* gene from such promoters will ensure that the inoculant strain survives in association with the host plant, but will senesce in the soil.

New methods for the identification and characterization of tagged introduced as well as resident rhizobial populations have been developed at Padova.

New biological material containing highly expressed reporter gene cassette, useful for monitoring purposes has been developed at Padova.

MAJOR COOPERATIVE LINKS

For Bristol, the main links within the project were with Heligenetics, Gaiba, and the University of Padua. These involved close research links and visits. Otherwise we developed close links with Gilson Manfio at Newcastle University who helped us to develop PYMS as a tool for these studies.

Work at Cork is continuing on the development of the containment system and monitoring of the field experiment. This work will progress under the IMPACT project in the BIOTECH programme. Strains will be reisolated from nodules of plants, grown in soil derived from the experimental site. Samples from these strains will be forwarded to Padua for fingerprinting. The field experiment will be effectively repeated in Italian soil by Heligenetics and extensively monitored by partners of this project.

PUBLICATIONS

Joint publictions

Corich V., Giacomini A., Ollero F.J., Squartini A. and Nuti M.P. (1991) Pulsed-field electrophoresis in contour-clamped homogeneous electric fields (CHEF) for the fingerprinting of *Rhizobium* spp. *FEMS Microbiol. Letters* **83**, 193-198.

Corich V., Bosco F., Giacomini A., Basaglia M., Squartini A. and Nuti M.P. (1993) The fate of genetically modified *Rhizobium leguminosarum* bv. *viciae* during inoculant manufacturing and storage. In 'Final sectoral meeting on biosafety and first sectoral meeting on microbial ecology', Granada, 24-27 october, *Abstr. Vol.,* pp.114-115.

Coutinho, H.L.C., Handley, B.A., Kay, H., Stevenson, L & Beringer, J.E. (1993). The effect of colony age on PCR fingerprinting. *Letters in Applied Microbiology* **17**, 282-284.

Kay, H.E., Fattori, M., Basaglia, M., Nuti, M.P. & Beringer, J.E.(1994) Factors influencing the survival of *Bradyrhizobium japonicum* in Italian soils. *Microbial Releases,* **140**, 2333-2339.

Kay, H.E., Coutinho, H.L.C., Fattori, M., Manfio, G.P., Goodacre, R., Nuti, M.P. & Beringer, J.E. (1994). The identification of *Bradyrhizobium japonicum* strains isolated from Italian soils. *Microbiology* **140**, 2333-2339.

Nuti M.P., O'Gara F., Basaglia M., Beringer J. (1993) Biosafety of genetically modified soil microbial inoculants. Commission of the European Communities, DGXII, BRIDGE Programme, European Laboratories without walls. Brussels, pp 1-11.

Squartini A., Dazzo F.B., Casella S., Nuti M.P. (1993) The root nodule symbiosis between *Rhizobium* 'hedysari' and its drought-tolerant host *Hedysarum coronarium. Symbiosis* **15**, 227-238.

Individual publications

Beringer, J.E., Kay, H.E. & Handley, B.A. (1994) Biosafety research: what is being done with bacteria? In 'Proceedings of the 6th European Congress on Biotechnology. *Eds. L. Alberghina, L. Frontali & P. Sensi, Elsevier, Amsterdam.*

Dowling, D., Boesten, B., O'Sullivan, D.J., Stephens, P., Morris, J. and O'Gara, F. (1992). Genetically modified plant-microbe interacting strains for potential release into the rhizosphere. In: *Pseudomonas*: Molecular Biology and Biotechnology (Eds. Galli, E., Silver, S. and Withold, B.) *ASM Publications, Washington D.C.* **44**, 408-414.
Giacomini A., Corich V., Ollero F.J., Squartini A. and Nuti M.P. (1992). Experimental conditions may affect reproducibility of the β-galactosidase assay. FEMS Microbiol. Letters **100**, 87-90.
Giacomini A., Ollero F.J., Squartini A. and Nuti M.P. (1994) Construction of multipurpose gene cartridges based on a novel synthetic promoter for high level gene expression in Gram-negative bacteria. Gene **144**, 17-24.
Nuti M.P., Squartini A., Giacomini A., Corich V., and Casella S. (1993) Molecular aspects of legumes/rhizobia symbiosis: perspectives for the '90s. Giornale Botanico Italiano **127**, 415-421.
Nuti M.P., Squartini A., Giacomini A. and Corich V. (1994) The use of genetically modified organisms (GMOs) in the environment: biosafety results of field tests and key scientific issues. In 'Proceedings of the 6th European Congress on Biotechnology' Florence, Italy, 13-17 June. L. Alberghina, L. Frontali, P. Sensi (Eds.), Vol. II, *Elsevier Science Publishers, Amsterdam*, pp. 1291-1296.
Nuti M.P., Squartini A., Giacomini A. (1994) European Community regulation for the use and release of genetically modified organisms (GMOs) in the environment. In 'Molecular Ecology of Rhizosphere Microorganisms' *F. O'Gara, D. Dowling, B.Boesten Eds., VCH Publ.* pp. 165-173.
O'Flaherty, S., Boesten, B. and O'Gara, F. (1994) A system based on thymidylate synthase, to improve plasmid maintenance in *Rhizobium* spp. under field conditions. *Microbial Releases* (submitted).
S. Roani, Roani M.C., Malagò C., Zuccolo E. and Basaglia M. (1992) Survival of *Bradyrhizobium japonicum* in soil is affected by water table level BRIDGE. *First sectoral meeting on Biosafety, Wageningen, The Netherlands,* Dec 6-9.
Wang, Y-P., Giblin, L., Boesten, B. and O'Gara, F. (1993) The Escherichia coli cAMP receptor protein (CRP) represses the *Rhizobium meliloti dct*A promoter in a cAMP-dependent fashion. *Molecular Microbiology* **8**, 253-259.
Zuccolo C., Roani M.C., Malagò C., Baldan S. and Basaglia M. (1993) The effect of water table adjustements on the survival of *Bradyrhizobium japonicum* in soil. *Final Sectorial Meeting on Biosafety and First Sectorial Meeting on Microbial Ecology, Granada,* October 24-27.

Methodology for the fast design of fungal DNA probes and PCR primers (BIOT CT-910301)

COORDINATOR:
(1) W. MOENS, Int. Hygiene & Epidemiology, Brussels, B
PARTICIPANTS
(2) J. FRISVAD, Danmark Tekniske Højskole, Lingby DK
(3) F. GANNON, Univ. College Galway, Galway, IRL
(4) L. ROSSEN, Biotechnological Institute, Lingby, DK
(5) V. RUBIO, Centro Nacionale de Biotecnologia, Madrid, E,
(6) N. NOLARD, IHEM Culture Collection, Brussels, B

BACKGROUND INFORMATION

The fungal world is estimated to be represented by 500-650.000 species of which 20% can be grown in culture and 10% have been identified. Filamentous fungi are usually classified on the basis of their morphology, enzymes, metabolites and, recently, their DNA and genetic properties.

Filamentous fungi (FF) being of pervasive importance in many areas, the availability of fungal DNA-based probes and molecular typing strategies should ultimately facilitate improved crop yields, greater productivity, increased safety for the consumer and the environment and improved quality control of industrial processes and products. They have major negative economical and health consequences as opportunistic pathogens, spoilage agents of stored crops and food. Equally, FF have diverse positive roles in the food and non food sectors and as biocontrol or production agents. These organisms are known to play key ecological roles in the environment and are responsible for many plant, animal and human diseases. As sporulators, pathogens and producers of antibiotics, toxic, mutagenic or carcinogenic metabolites, natural and genetically-modified fungi raise important environmental and health questions of biosafety, agro/food management and quality. For these reasons, the reliable identification of the fungi and a well developed taxonomy were -and still are- an urgent need in the EC.

DNA probes for FF generated in a sound phylogenetic context using rigourous standardised DNA technology were not only needed for agro-industrial applications but also for fungal epidemiology, biodiversity studies and the monitoring of FF of biotechnology interest.

OBJECTIVES AND PRIMARY APPROACHES

The use of DNA probes for the rapid detection of specific microorganisms is becoming established in clinical diagnostics, the food industry and environmental analysis. Specific probes have now been developed for a variety of organisms. Many of these probes have been based on DNA sequences from the ribosomal gene complex. These DNA sequences afford particular advantages as targets for the development of DNA probes. Firstly, they are present in multiple copies/ genome. Secondly they contain highly conserved sequences which can be used as common targets for PCR consensus primers as shown by 3. Thirdly, they contain highly diverse spacer regions which in some cases can be exploited for the development of species-specific probes. To date procaryotic pathogens have been the primary targets for DNA probe analysis.

To solve the major bottleneck defined by the general lack of fungal genus and species-specific DNA probes or PCR markers, the participants have been working along three lines:

(i) combination of DNA amplification techniques (RAPD and PCR), DNA sequencing, sequence analysis and taxonomic definition of fungal ecological niches,

(ii) complementation between mycologists and molecular biologists, and

(iii) harmonisation of the basic methodology between the involved participating laboratories.

Two main objectives were defined:

(i) Identification of a region of the ribosomal genes amenable to genus and/or speciesprobe design after amplification, sequencing and sequence alignment,

(ii) Exploration of the molecular typing potential of amplified DNA polymorphism for taxonomical pre-screening and strain certification.

Two types of genomic targets were selected at start:

(i) The ubiquitous ribosomal DNA sequences were chosen as fungal genomic sequences of common interest. The fungal ribosomal operon can be subdivided into 5 majors parts: the intergenic sequence (IGS) nesting the 5S subunit, the 18S subunit, the internally transcribed sequence (ITS) nesting the 5.8S sub-unit, and the 28S sub-unit. The transcribed sequences (5S, 18S, 5.8S, 28S) are highly conserved between genera, family and even phyla. As in the case of procaryotic spacers, the fungal IGS and ITS regions were expected - and further demonstrated by the participants in the case of the fungal ITS region- to be highly variable.

(ii) DNA polymorphisms as generated by arbitrarily-primed PCR using fungal universal consensus ribosomal primers as single amplimer .

RESULTS AND DISCUSSION

Several DNA techniques were adapted to fungi and standardized among participants using representative fungal genera and species involved in several areas of biotechnology. (*Aspergillus* sp., *Cryptococcus* sp., *Mucor* sp., *Penicillium* sp., *Trichoderma* sp. *Rhizoctonia* sp.). Scientists were exchanged and a strong quality control structure was used for strain certification, primers and buffer distribution. A DNA bank from certified strains officially deposited in the 1 -linked fungal IHEM Culture Collection has been established.

As a result, a DNA technology and scientific basis were made operational. All steps from strain collection to species-specific PCR primer validation including PCR tagging, DNA sequencing and sequence analysis, definition of ecological niches have been analysed by one or more of the members. In parallel, participants 1, 3,4 have developed different fungal DNA probes using similar standardized techniques.

Two protocols of direct sequencing of amplified ribosomal 18S fragment were carried out in 3 and 4 lab. Several fast DNA extraction (1,3,4) and PCR product direct sequencing (3,4) protocols were evaluated. PCR amplification conditions of the ribosomal 18S and ITS regions were generalised to 97 species covering 24 genera (1,3,4,5,6). Participant 2 and 6 participated to all studies as an reference taxonomist. Results can be sub-divised into four parts:

(1) Quick DNA preparation methods

A first protocol was set up in 4. It is based on disruption of spores and mycelium by freeze-thawing without the use of enzymes followed by phenol-chloroform extraction and DNA precipitation. It was found convenient by all participants for small cultures of most fungi tested including those with very resistant cell walls. The resulting DNA could be used for PCR and direct sequencing of PCR products. A second protocol was designed by 1 which is based on glass beads disruption of spores and mycelium in the presence of phenol and SDS followed by DNA precipitation. The two protocols were suitable for both PCR, RAPD and direct DNA sequencing of PCR products.

(2) Assessment of fungal nuclear consensus PCR primers (NS and ITS series)

The universality of the consensus primers for the 18S ribosomal subunit (NS primers) and for the ITS region (ITS primers) was assessed by all participants on the DNA's isolated from 97 species representative of *Absidia, Acremonium, Aspergillus, Candida, Cladosporium, Cryptococcus, Epidermophyton, Fusarium, Geotrichum, Microsporum, Mortierella, Mucor, Paecilomyces, Penicillium, Phialosphora, Pityrosporum, Rhizoctonia., Rhizomucor, Rhizopus, Scopulariopsis, Synecephalastrum, Trichoderma, Trichophyton and Trichosporon.*

A common PCR format was validated by all participants. This format was suitable for all species tested including those producing PCR inhibitors or whose ribosomal DNA sequence conformation interfered with PCR and sequencing reactions. Primers of the NS series were working with all fungi tested whereas only ITS5 and ITS4 primers could be defined as universal.

(3) Analysis of the DNA sequences of the fungal ribosomal genes

Several parts of the fungal ribosomal genes from various fungal genera, species, varieties and types have been directly sequenced and analysed after DNA amplification using consensus primers. The data generated by 3 and 4 showed that the 18S region was not a rich source of potential DNA probe or species specific PCR primer sequences. Therefore, ITS DNA sequences from 54 species covering 164 isolates were generated by participants 1, 3, 4, 5. From these data, the two internally transcribed ITS sequences adjacent to the 5.8S sub-unit appeared as regions of enormous potential for family-, genus- and species-specific probes and the study of mold phylogeny. Accordingly, probes have been designed from these data by the same participants. The discrimination potential between fungi appeared to be limited at the species level: in 2 pilot studies carried out by participants 1 on *Aspergillus fumigatus* and by 4 on *Penicillium aurantiogriseum*, a very high degree of ITS DNA sequence homology was found at the intra-species level.

(4) PCR analysis of intra-species genotype diversity

Fungal sub-species genotypes were characterized at multiple genetic loci by 1 on basis of the classes of DNA Polymorphisms revealed by DNA amplification using single ribosomal consensus primers from the NS and ITS series. Large strain collections of the *Aspergillus fumigatus* and *Cryptococcus neoformans* were analyzed with different primers. The discrimination potential at each locus was calculated and found independent of the GC genomic content. Combination of the multilocus data increased the discrimination potential of the method up to, or above the one of Restriction Fragment Length Polymorphims analysis.

Most importantly, a rationale linking genotypes, genotype population frequencies and loci was designed allowing to build maps of intra-species genetic diversity.Further DNA sequence analysis of DNA polymorphisms should provide a molecular basis to this new mapping system. The methodology has already reached the prenormative stage since:

(i) Analysis of fungal DNA polymorphisms was developped under U.E. GLP's;

(ii) a set of *A. fumigatus* strains defined by this DNA polymorphism-genotyping approach has been choosen as common reference biological material by a EORTC network of European mycopathologists.

In conclusion, genotyping DNA polymorphism's and PCR probing were complementary since their combined taxonomical discrimination potentials ranged between the family down to the strain-specific genotypes. Such complementarity is documented by a joint report from 4, 2 and 1 and individual reports from 1.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The concerted action provided an opportunity to exploit a coherent methodology based on the standardised combination of classical and the last DNA amplification techniques for the following purposes:

(1) Inter- and intra-species- genetic diversity at multiple loci
(Applications: taxonomical pre-screening computerized expert system, population diversity and structure, process quality control, Culture collection management, strain certification)

(2) Phylogeny based on the DNA sequencing of the ITS regions of the fungal ribosomal genes
(Applications: Phylogenetics, Taxonomy, fast identification and validation of ribosomal genes-based DNA probes, probing fungi by PCR, sandwich-PCR, LCR or flow cytometry.

DNA probes are destinated to agro-food-fish research and industries, medical and veterinean research and clinical diagnostics, biosafety and intellectual property protection).
Although the methodology was specifically adapted to fungi, similar approaches are reported for animals, plants, insects, bacteria and viruses.

MAJOR COOPERATIVE LINKS

Contact group meetings:
June 92-Madrid, September 92-Brussels, November 92-Copenhagen, Wageningen 92, June 93-Galway, November 93, Granada .
Exchange of certified material: fungal strains and DNA batches, plasmids, DNA sequences.
One week DNA sequencing training course in 4 lab for one 1 and one 5 participant.

Internal collaborations:
5 + 4 for the sequencing of *Rhizoctonia* spp., 1+2+4 for a taxonomical study about correlations between PCR fingerprinting and chemotyping of *Penicillium* spp., B, DK, E, I for the characterization of the fungal ITS ribosomal region, 5 and 1 for the typing of dsRNA-containing *Aspergillus* sp.

PUBLICATIONS

Joint publications

W. Moens, F. Gannon, N. Nolard, L. Rossen, V. Rubio (1992). Fast design of fungal PCR markers.Abstract of the 1st European Conference on Fungal Genetics, Nottingham 20-23/10/92

W. Moens, J. Frisvad, F. Gannon, N. Nolard, L. Rossen, V. Rubio (1992). Identification and taxonomy analysis of molds by use of oligonucleotide probing EC-BRIDGE Biosafety sectorial meeting (1992)

W. Moens, C. Dawson, B.Van Vaerenbergh, H. Beguin, P. Skouboe, V. Colombo, Ines Canosa, M. Dawson, F. Gannon, N. Nolard, L. Rossen, V. Rubio (1992). Methodological standardisation between four molecular biology laboratories and the IHEM Culture Collection: taxonomic validation of fungal ribosomal 18S and ITS consensus primers as amplimers for the PCR amplification of various parts of the fungal ribosomal operon. EC-BRIDGE Biosafety sectorial meeting

W. Moens, C. Dawson, B.Van Vaerenbergh, H. Beguin, P. Skouboe, V. Colombo, Ines Canosa, M. Dawson, F. Gannon, N. Nolard, L. Rossen, V. Rubio (1992). Taxonomic validation of fungal ribosomal 18S and ITS consensus primers as amplimers for the PCR amplification of various parts of the fungal ribosomal operon. Communication at the French Society of Mycology (Paris 27/11/92)

B. Van Vaerenbergh and W. Moens, Pernille Skouboe and Lone Rossen(1992). Amplification Fragment Length Polymorphisms of *Aspergillus fumigatus* strains revealed after DNA amplification using fungal ribosomal 18S consensus primers as single amplimers. EC-BRIDGE Biosafety sectorial meeting

M.T. Dawson, P. Skouboe, C. Dawson, W. Moens, L. Rossen and F. Gannon (1992). A comparison of the ribosomal sequences from *Aspergillus niger* and *A. fumigatus*. EC-BRIDGE Biosafety sectorial meeting

M.T. Dawson, N. Nolard, W. Moens and F. Gannon (1992). Analysis of the ITS ribosomal spacer region of the genus Mucor and related species: a site for DNA probe development (1992) EC-BRIDGE Biosafety sectorial meeting

Dawson et al (1, 3,4,5) (1994) Characterization of the fungal ribosomal ITS region as a source of DNA probes.(in preparation)

Frisvad et al (1,2,4) (1994) Characterization of *P. aurantiogriseum* taxonomy using chemotaxonomy, DNA fingerprinting and DNA-based phylogenetics (in preparation).

Skouboe et al (1,3,4,5) (1994) Validation of *Penicillium*-specific probes derived from the ITS ribosomal region. (in press).

Van Vaerenbergh et al (1,2,3,4,5) PCR Amplification of the fungal ITS ribosomal region: universality status of established and new consensus primers (1994) (in preparation).

Individual publications

B. Van Vaerenbergh, H. Beguin, N. Nolard and W. Moens(1992) Amplification Fragment Length Polymorphisms of *Aspergillus fumigatus* strains reveals population heterogeneity and allows potential identification at the strain level. Meeting of the French Society of Mycology

B. Van Vaerenbergh, N. Nolard and W. Moens (1992) Ribosomal operon-derived probes for the PCR identification of opportunistic mycopathogens. Meeting of the French Society of Mycology and EC-BRIDGE-Biosafety

W. Moens (1993) Fingerprinting and probing *Cryptococcus neoformans*-2nd International Conference on *Cryptococcus* and Cryptococcosis

B. Van Vaerenbergh, B. Grootaert and W. Moens (1994) Validation of a method for the preparation of fungal DNA suitable for PCR and RAPD, J. Med. Mycology (in press)

B. Van Vaerenbergh, F. Symoens, N. Nolard and W. Moens (1994) Genotyping of *A. fumigatus* intra-species diversity at multiple loci by RAPD (submitted)

B. Van Vaerenbergh, B. Grootaert, D. Verbruggen, H. Beguin, D. Swinne, N. Nolard and W. Moens (1994) Genotyping *C. neoformans* intra-varieties diversity at multiple loci by RAPD (submitted)

BRIDGE

T-PROJECTS

T-PROJECT

'SEQUENCING OF THE YEAST GENOME'

Sequencing of the yeast genome (BIOT CT-900167)

COORDINATOR:
Université Catholique de Louvain, Louvain-La-Neuve, B

Chr II

COORDINATOR:
H. FELDMANN, Universität München, München, D
PARTICIPANTS:
M. AIGLE, Université de Bordeaux, Bordeaux, F
C. CHRISTIANSEN, Genetic Engineering Group, Lyngby, DK
H. DOMDEY, Genzentrum, München, D
K.D. ENTIAN, Universität Frankfurt, Frankfurt, D
H. FELDMANN, Universität München, München, D
W. FIERS, Rijksuniversiteit Gent, Gent, B
F. FOURY, U.C.L., Louvain-La-Neuve, B.
N. GLANSDORFF, Res. Inst. CERIA-COOVI, Bruxelles, B
L. GRIVELL, Universiteit van Amsterdam, Amsterdam, NL
C. HOLLENBERG, Universität Düsseldorf, Düsseldorf, D
C. JACQ, Ecole Normale Supérieur, Paris, F
M. JACQUET, Université de Paris-Sud, Orsay, F
C. JAUNIAUX, DKFZ, Heidelberg, D
P. NETTER, CNRS, Gif-sur-Yvette, F
F. POHL, Universität Konstanz, Konstanz, D
M. RIEGER, Biot. Bio. Forsch., Wilhelmsfeld, D
P. SLONIMSKI, CNRS, Gif-sur-Yvette, F
J.L. SOUCIET, CNRS, Strasbourg, F
H.Y. STEENSMA, University of Leiden, Leiden, NL
K. WOLFE, University of Dublin, Dublin, IRL

Chr XI

COORDINATOR:
B. DUJON, Institut Pasteur, Paris, F
PARTICIPANTS:
W. ANSORGE, EMBL, Heidelberg, D
B. DUJON, Institut Pasteur, Paris, F
H. DOMDEY, Genzentrum, München, D
K.D. ENTIAN, Universität Frankfurt, Frankfurt, D
F. FOURY, U.C.L., Louvain-La-Neuve, B.
H. FUKUHARA, Institut Curie, Orsay, F
F. HILGER, Fac. Sc. Agronomique, Gembloux, B
C. HOLLENBERG, Universität Düsseldorf, Düsseldorf, D
C. JAUNIAUX, DKFZ, Heidelberg, D
A. JIMENEZ, Universidad Autonoma, Madrid, E
S. OLIVER, University of Manchester, Manchester, UK
P. PHILIPPSEN, Universität Giessen, Giessen, D
R.J. PLANTA, Vrije Universiteit, Amsterdam, NL
F. POHL, Universität Konstanz, Konstanz, D
M. RIEGER, Biot. Bio. Forsch., Wilhelmsfeld, D
C. RODRIGUES, Instituto Gulbenkian, Oeiras, P

G. THIREOS, IMBB, Iraklio, GR
D. VON WETTSTEIN, Carlsberg Lab., Kobenhavn, DK

OBJECTIVES AND PRIMARY APPROACHES

General administration and coordination of the project

- Subcontracts to DNA coordinators (Chromosomes II, VII, X, XI and XIV) and 36 sequencing laboratories.
- Initial advance payments to each subcontractor.
- Detailed accountancy of the base pairs produced by each sequencing laboratory.
- Subsequent payments on the basis of 2 ECU per final base pair sequenced as approved by the DNA and informatic coordinators.
- Subsequent payments of the DNA coordinators according to progress.
- Continuous interaction with each DNA coordinator, the informatic coordinator and the sequencing laboratories.
- Organisation of contractors meetings.

Preparation of organised libraries from chromosome X (Galibert), XIV (Philippsen) and VIII (Oliver)
Complete sequencing of chromosome XI (660 kb) and partial sequencing of chromosome II (800 kb)

- Construction of two libraries of overlapping clones to cover the entire yeast chromosome II and XI as unique contigs.
- Sorting of clones and construction of high resolution ('sequence-ready') physical maps.
- Distribution of clones to participants for sequencing until the entire chromosomes are finished.
- Application of 'quality controls' to the sequences determined by the participants.
- Assembly of the complete sequence of chromosome II and XI from the sets of overlapping clones and interpretation of the sequence.

Organisation of contractors meetings in Munich (October 1992), Louvain-la-Neuve (April 1993) and Manchester (February 94).

RESULTS AND DISCUSSION

Chr II

The complete DNA sequence of the yeast *Saccharomyces cerevisiae* chromosome II has been determinated and submitted for publication in 1994. This is the largest eukaryotic chromosome entirely sequenced (807,188 bp). A total of 410 open reading frames (ORFs) were identified, covering 72% of the sequence. Similarity searches revealed that 121 ORFs (30%) correspond to genes of known function, 51 ORFs (12.5%) have homologues among gene products from yeast or other organisms whose functions are known, while 52 others (12.5%) have homologues the functions of which are not well-defined and another 33 of the novel putative genes (8%) show levels of similarity of uncertain significance. 37 to 45% of the genes on chromosome II are thus of unpredicted function. Among the novel putative genes, we found several that are related to genes that perform differentiated functions in multicellular organisms or are involved in malignancy. Beyond the many novel genes detected in chromosome II, the analysis of this chromosome substantiated general chromosome patterns but has also revealed particular novel features of chromosomal

organisation. Alternating regional variations in average base composition correlating with variations in local gene density along a chromosome were confirmed. We propose that putative functional ARS elements are preferably located in the (A+T) rich regions that have a spacing of ca 110 kb. In chromosome II, the distribution of coding sequences between the two strands is biased, with a ratio of 1.3. An interesting aspect of eukaryotic genome organisation and evolution is the finding that chromosome II has a high degree of internal genetic redundancy.

Chr XI

The complete DNA sequence (666, 448 base pairs) of the yeast *Saccharomyces cerevisiae* chromosome XI has been determinated and published. In addition to a compact arrangement of potential protein coding sequences, the 666,448 base pairs sequence has revealed general chromosome patterns; in particular, alternating regional variations in average base composition correlate with variations in local gene density along the chromosome. Significant discrepancies with the previously published genetic map demonstrate the need for using independent physical mapping criteria. A total of 331 ORFs were identified in the entire chromosome. Seven of these ORFs are interrupted by introns. Comparison of the present sequence with public databases revealed that 93 of the 331 ORFs (28%) correspond either to previously known protein-encoding genes or to genes whose functions have been determined during this work. All other ORFs, 72% of the total, represent novel putative yeast genes whose functions need to be experimentally determined. But 93 of them (another 28% of the total) have homologues among gene products from yeast or other organisms whose functions are known, whereas 37 others (11% of the total) have homologues that are themselves of unknown function. The remaining 108 ORFs (33% of the total) either have no homologues in data libraries or show levels of similarity of uncertain significance. Overall, about 40-44% of the genes of chromosome XI are thus of unpredicted function, a figure similar to that of chromosome III.

The very high gene density previously found with chromosome III is confirmed: ORFs occupy on average 72% of the sequence of chromosome XI. The average ORF size is 488 codons (1,464 bp).

A higher level of accuracy (99,97%) has been achieved in this project only after costly effort; it still implies that about one-third of predicted genes will contain sequencing errors that will affect their interpretation.

Highlights/Milestones

Year:	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*
Sequences submitted by the EU Network (kb)	115	212	215	876	1239	>2000
Cumulative (kb):	115	327	532	1408	2647	>4700

Wider considerations

Among all eukaryotic model organisms, *Saccharomyces cerevisiae* combines several advantages:

(i) this yeast has a genome size of only 13.5 megabases, i.e. 200 times smaller than that of the human genome;
(ii) the yeast system is susceptible to powerful genetic techniques, and
(iii) functions in yeast have been studied in great detail by biochemical approaches.

Based on present data, one can calculate that a repertoire of some 6500-7000 genes is sufficient to build this simple eukariotic cell. Considering recent progress and a world wide enterprise of sequencing the yeast genome (see below), we can be confident to decipher its genetic potential within a reasonable time period and with a relatively limited effort.

Since a large variety of examples provide evidence that substantial cellular functions are highly conserved from yeast to mammals and in many cases corresponding genes can complement each other, the wealth of sequence information obtained in yeast will be extremely useful as a reference against which sequences of human, animal or plant genes may be compared. Moreover, the ease of genetic manipulation in yeast opens the possibility to functionally dissect novel genes from other eukaryotes in the yeast system.

The rationale behind our effort on quality control was that if systematic genome sequencing of yeast is to be useful and significant, then sequence accuracy must permit correct interpretations. With the gene density and ORF size distribution of yeast, even relatively rare sequencing errors result in a large fraction of the protein-coding genes being affected by frame shifts. A simple calculation predicts that with an accuracy of 99% virtually all predicted genes contain errors, and that at 99.9% accuracy, two-thirds of the genes still contain at least one error.

Two years ago, our consortium of 35 European laboratories published the first complete sequence of a eukaryotic chromosome, that of chromosome III of *S. cerevisiae* (*Nature* **357**, 38-56). During the past three years, our consortium has turned its efforts to the sequencing of yeast chromosomes XI and II and will continue to contribute to the sequencing of the yeast genome.

The contributions of the 11 different European countries (31 laboratories) to the achievement of the BRIDGE sequencing task are shown in the figure 1.

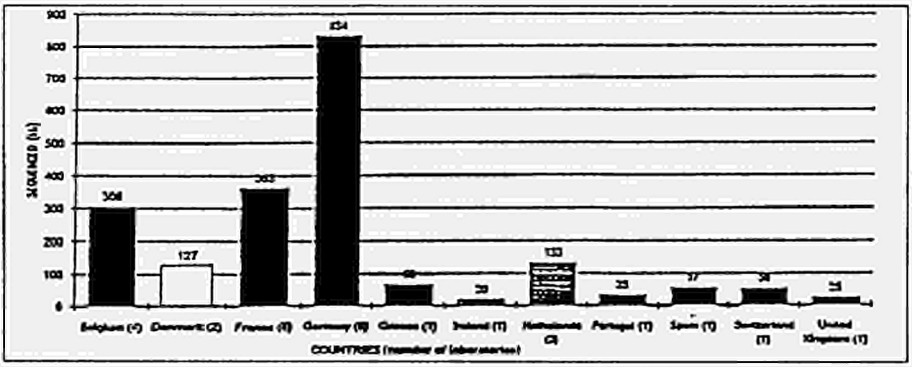


Fig. 1: Contribution of each country to the BRIDGE program

MAJOR SCIENTIFIC BREAKTHROUGHS

Europe is the undisputed leader in this field.
The cumulative capacity of our European network is shown in the figure 2.

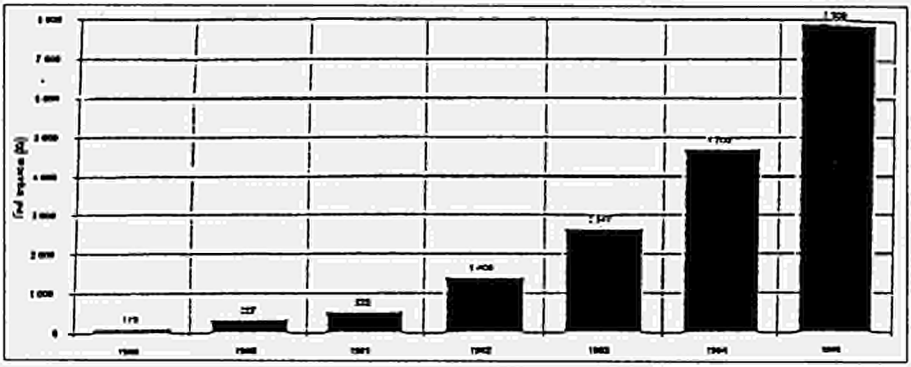


Fig. 2: Evolution of the yeast genome sequencing cumulative capacity of the EU.

The efficiency of the yeast genome sequencing network has increased appreciably. An increasing flux of 1.2 Mb/year final sequence was attained in 1993 corresponding to more than one complete gene per day.

MAJOR COOPERATIVE LINKS

The 36 European laboratories participating to this work have met in Munich (October 1992), Louvain-la-Neuve (April 1993) and Manchester (February 1994).

A world wide agreement has been reached to distribute among EU, USA, Canada, Japan and Cambridge (UK) the sequencing of the 16 yeast chromosomes (Goffeau, A. Nature 369, 101-102 (1994)). A calendar has been agreed upon. If all parties respect their schedule, the entire sequence of the entire yeast genome should be terminated in 1996.

Yeast genome sequencing world wide agreement (funding secured — July 1994)

CHR	*KB*		*Coordinator*	*Affiliation*	*Release date*
I	235	CAN	H. Bussey	McGill Univ., Montreal	94
II	807	EU	H. Feldmann	Munich Univ.	94
III	314	EU	S. Oliver	UMIST, Manchester	done
IV right	500	USA	R. Davis	Stanford Univ.	95
IV middle	540	UK	B. Barrell	Sanger C., Cambridge	95
IV left	600	EU	C. Jacq	ENS, Paris	96
V	550	USA	R. Davis	Stanford Univ.	94
VI	270	JAP	Y. Murakami	Riken, Tsukuba	95
VII	1150	EU	H. Tettelin	Louvain-La-Neuve Univ.	96
VIII	550	USA	M. Johnston	Washington Univ., St Louis	94
IX	440	UK	B. Barrell	Sanger C., Cambridge	94
X	720	EU	F. Galibert	CNRS, Rennes	95
XI	666	EU	B. Dujon	Pasteur, Paris	done
XII right	750	USA	M. Johnston	Washington Univ., St Louis	95
XII left	450	EU	J. Hoheisel	DKFZ, Heidelberg	96
XIII	910	UK	B. Barrell	Sanger C., Cambridge	95
XIV	810	EU	P. Philippsen	Basel Univ	95
XV	1150	EU	B. Dujon	Pasteur, Paris	96
XVI right	450	UK	B. Barrell	Sanger C., Cambridge	96
XVI left	592	CAN	H. Bussey	McGill Univ., Montreal	96

PUBLICATIONS

Joint publications
Dujon, B. et al. Nature **369**, 371-378 (1994)
Feldmann, H. et al. EMBO, (submitted)

Individual publications
Chr II
Baur, A. *et al. Yeast* **9**, 289-293 (1993).
Bécam, A.M. *et al. Yeast* **10**, S1-S11 (1994).
Bussereau, F. *et al. Yeast* **9**, 797-806 (1993).
Delaveau, T. *et al. Yeast* **8**, 761-768 (1992).
Démolis, N. *et al. Yeast* **9**, 645-659 (1993).
Doignon, F. *et al. Yeast* **9**, 189-199 (1993).
Doignon, F. *et al. Yeast* **9**, 1131-1137 (1993).
Holmstrom, K. *et al. Yeast* **10**, S47-S62 (1994).
Mallet, L. *et al. Yeast* **10**, 819-831 (1994).
Miosga, T. *et al. Yeast* **9**, 1273-1277 (1993).
Schaaff-Gerstenschläger, I *et al. Yeast* **9**, 915-921 (1993).
Scherens, B. *et al. Yeast* **9**, 1355-1371 (1993).
Skala, J. *et al. Yeast* **8**, 777-785 (1992).
Skala, J. *et al. Yeast* **10**, S13-S24 (1994).

Smits, P. *et al. Yeast* **10**, S75-S80 (1994).
Van Dyck, L. *et al. Yeast* **8**, 769-776 (1992).
Wolfe, K. *et al. Yeast* **10**, S41-S46 (1994).

Chr XI

Alexandraki, D. *et al. Yeast* **10**, S81-S91 (1994).
Bossier, P. *et al. Yeast* **10**, 681-686 (1994).
Bou, G. *et al. Yeast* **9**, 1349-1354 (1993).
Boyer, J. *et al. Yeast* **9**, 279-287 (1993).
Chéret, G. *et al. Yeast* **9**, 1259-1265 (1993).
Colleaux, L. *et al. Yeast* **8**, 325-336 (1992).
Düsterhöft, A. *et al. Yeast* **8**, 749-759 (1992)
Garcia-Cantalejo, J. *et al. Yeast* **10**, 231-245 (1994).
Jacquier, A. *et al. Yeast* **8**, 121-132, (1992).
James, C. *et al. Yeast* **10**, 257-264 (1994).
James, C. *et al. Yeast* **10**, 247-255 (1994).
Pallier, C. *et al. Yeast* **9**, 1149-1155 (1993).
Pascolo, S. *et al. Yeast* **8**, 987-995 (1992).
Purnelle, B. *et al. Yeast* **8**, 977-986 (1992).
Purnelle, B. *et al. Yeast* **10**, 125-130 (1994).
Purnelle, B. *et al. Yeast* **9**, 1379-1384 (1993).
Thierry, A. et al. Nucleic Acids Res. **20**, 5625-5631 (1992).
Rasmussen, S.W. Yeast **10**, *S63-S68 (1994).*
Rasmussen, S.W. Yeast **10**, *S69-S74 (1994).*
Singer-Krüger, B. et al. J. Cell Biol. (in press).
Thierry, A. et al. Nucleic Acids Res. **20**, 5625-5631 (1992).
Tzermia, M. *Yeast* **10**, 663-679 (1994).
van Vliet-Reedijk, J.C. *et al. Yeast* **9**, 1139-1147 (1993).
Vandenbol, M. *et al. Yeast* **10**, S25-S33 (1994).
Vandenbol, M. *et al. Yeast* **10**, S35-S40 (1994).
Wiemann, S. *et al. Yeast* **9**, 1343-1348 (1993).
Forrova, H. *et al. Yeast* **8**, 419-422 (1992).

Informatics Network (Sequencing the Yeast Genome) (BIOT CT-900161)

COORDINATOR:
H.W. MEWES, Max-Planck-Institut für Biochemie, Martinsried, D

BACKGROUND INFORMATION

Coordinated sequencing of the yeast genome by a network of independent laboratories requires centralized support for the distribution of genetic material and for the collection, processing, and evaluation of the experimentally determined data. After successful completion of chromosome III in 1992, efforts have been directed toward the sequencing of chromosomes II (ca. 820 kb) and XI (666 kb). In close collaboration with the DNA Coordinators, the Informatics Coordination Center provided the participating laboratories with support for sequence verification, quality control, and sequence data analysis.

OBJECTIVES AND PRIMARY APPROACHES

As the informatics coordinator, MIPS has been responsible for the accumulation, correlation, storage, and analysis of the sequence data generated by the 33 laboratories participating in the chromosomes II and XI (820 and 666 kb, respectively) sequencing projects.

RESULTS AND DISCUSSION

As of December 1993, 1.4 Mbases of non-redundant DNA have been submitted to MIPS. The sequence of chromosome XI has been completed; a publication describing this work has been submitted. The sequence of chromosome II is expected to be published by mid 1994.

The yeast chromosome sequencing projects follow a network model involving a set of DNA-coordinators, the participating laboratories, and an informatics coordination center. The DNA-coordinators are responsible for selecting and distributing clones from an ordered library based on precise physical maps. The laboratories perform the experimental sequencing work. MIPS serves as the informatics coordinator and is responsible for data processing and chromosome assembly.

The work at MIPS has been performed in close collaboration with the DNA coordinators (B. Dujon, Paris and H. Feldmann, Munich) to ensure the highest level of quality control. Sequences submitted to MIPS have been subjected to systematic analysis of coding and intergenic regions. The analyses have been performed using the PIR-NBRF Query System (ATLAS) and the Genetics Computer Group (GCG) Sequence Analysis Software, and several additional programs for sensitive sequence comparison obtained from the EMBL file server. Detailed plots displaying the coverage of sequenced clones, all open reading frames and previously mapped genes, and other elements (Ty, LTRs, tRNA) were constructed. All sequences have been inspected for consistency with the physical map, for frameshift anomalies, and for consistency among overlapping sequences. Homology searches were carried out at the nucleotide level in order to detect possible vector contaminations, non-coding regions of significance, and regulatory sequences like splicing signals, promoters, and enhancers.

A total of 770 open reading frames have been identified and analyzed. Of the 330 chromosome XI open reading frames, 92 correspond with known yeast genes (28%). 168 (51%) show significant homology with previously published sequences while 122 (37%) do not reveal any significant similarity; the remainder display low levels of similarity and must be subjected to more detailed analysis before postulating potential functions.

Laboratories participating in the directed sequencing of chromosome XI have independently published their results and submitted 23 entries to the public nucleic acid sequence databases, totaling 309,467 bases with an average length of 13,455 bases/submission. This contrasts with 107 chromosome XI entries, generated exclusive of the yeast genome sequencing project, totaling 288,252 bases with a average length of 2,694 bases/submission. The genome project has contributed significantly more complete, larger chromosomal segments of significantly greater quality and accuracy than those generated exclusive of this initiative.

Three independent methods were applied to estimate the accuracy of sequencing within this project: comparison of data obtained from overlapping sequenced regions; controlled sequencing of randomly selected clones; and directed resequencing. The average error rate was determined to be about 0.03%, i.e. 3/10.000 bases. 157 kb of the final sequence were verified in total.

Comparison of the data generated in this project with the corresponding sequences in the public nucleic acid sequence databases also allowed estimates to be made concerning the quality of sequences deposited in these public resources. The error rate determined in this study was used to calibrate the number of discrepancies actually observed. The estimated error rate in public databases is 3-5 times higher than that observed in chromosome XI project.

Fig. 1: Graphical representation of sequencing results (XCHROMO)

A database of sequence similarities (FASTA-database) has been developed to allow the similarity of previously unidentified yeast sequences with new sequence data to be continuously monitored. All protein sequences collected by the PIR-International group are inserted routinely in the database allowing instantaneous access to the latest sequence similarity results. Yet unpublished, confidential sequences are also included in the database, but are accessible only by authorized members of the MIPS group.

A program for the visual inspection of yeast sequencing data has been developed (XCHROMO). It provides a graphic display of the current status of yeast chromosome assembly and displays important landmarks, including the locations of cosmids, restriction sites, open reading frames, and genetic elements. Fig. 1 shows an example depicting data from chromosome X.
In collaboration with E. Sonnhammer (Cambridge) existing software (ACEDB), developed for the *Caenorhabditis elegans* project has been employed to provide a suitable medium for data distribution and to provide an intelligent interface to the data. The system allows access to homologous sequences as found by the BLAST sequence homology search program.

On-line computing facilities are provided by MIPS to enable the participating laboratories to analyze their data and perform database queries. Regular summaries of progress on each chromosome are provided to the contributing laboratories and the yeast industry platform (YIP).

Summary of sequence data processing

All new sequence data submitted to MIPS are compared with the previous submissions of the same contractor, overlaps are identified, and contigs are assembled accordingly. Restriction fragment data of the resulting contigs are compared with the physical maps provided by the DNA coordinator for consistency. Subsequently the contigs are compared to the adjacent sequences to detect overlaps for the stepwise assembly of the entire chromosome.
On the DNA level systematic searches are performed to detect:
— known yeast sequences
— tRNA genes
— LTRs
— regulatory elements (UAS, ARS, promoters, repeats, etc.)
— introns

Open reading frames (ORFs) are extracted with a cutoff value of 100 amino acids and the surrounding sequence is examined for the presence of potential regulatory elements. The codon usage and codon adaptation index are computed for each ORF and their probability of expression is assessed.

Hypothetical proteins translated from ORFs are further analyzed by similarity comparison against the most recent version of the protein sequence database. The sequences are partitioned into three classes: strongly homologous, possibly related, and unrelated proteins. Strong homologs may correspond to previously published yeast proteins with known function or to non-yeast homolog proteins; the functions of these proteins can be assigned unambiguously and they can be classified by protein family. Pattern searches are performed against the ProSite Dictionary of Protein Sites and Patterns (compiled by A. Bairoch, Geneva). The ORFs are also inspected for internal repeats, putative transmembrane segments, etc. Local similarity to other evolutionarily-related proteins is analyzed with the BLAST

program (Altschul et al.). The results are stored in the form of a database allowing dynamic retrieval of sequence analysis data.

MAJOR SCIENTIFIC BREAKTHROUGHS

25% of the yeast genome (ca. 14 million base pairs) has been elucidated and published. During the current contract, the sequences of the first two chromosomes (III and XI) have been experimentally determined by the laboratories of the consortium and processed and analyzed by the informatics coordinator. More than 1000 potential protein sequences have been discovered by the EU yeast sequencing program. About half of these do not correspond with any previously determined protein sequences and more than 60% can not be functionally classified using sequence comparison methods. This work has spawned a large number of more detailed investigations, both experimental and theoretical.

Work is currently underway to produce a series of CD-ROMs dedicated to the yeast genome projects. The CD-ROMs will be produced as the data accumulates and will contain the latest collections of known yeast chromosomal sequences and associated information (e.g., protein sequences, homologous sequences, annotation, etc.). Software will be included to provide access to these data. The CD-ROM will provide an excellent vehicle to make the results of the yeast chromosome sequencing projects available to the general scientific research community.

MAJOR COOPERATIVE LINKS

The sequence of chromosome XI has been completely processed and analyzed (666 kb, 330 identified open reading frames). 799 kb have been received from chromosome II. Results including matches within the confidential database have been forwarded to the laboratories. With the help of careful sequence analysis, sequencing errors have been corrected and a number of problems have been resolved.

In collaboration with the protein sequence databank group, MIPS prepares annotated, merged protein sequence databank entries for all ORFs found in the chromosomal DNA sequence. These entries are integrated into the protein sequence database at the time of publication.

Providing coordination and informatics support for a large-scale distributed project such as the European yeast sequencing initiative is a highly interactive process and demands frequent communication among all participants. In addition to employing state-of-the-art telecommunications, meetings with the participating laboratories were hosted, and frequent visits were arranged with the DNA-coordinators. The *Arabidopsis thaliana* sequencing project has been initiated in cooperation with the yeast projects, with MIPS serving also as the informatics coordinator.

PUBLICATIONS

A detailed report for the Yeast Industry Platform is issued quarterly. Chromosome Newsletters were distributed to the contractors in collaboration with the DNA-coordinators. The final complete sequence of chromosome XI has been submitted for publication.

T-PROJECT

'MOLECULAR IDENTIFICATION OF NEW PLANT GENES'

Molecular identification of new plant genes (BIOT CT-900207)

COORDINATOR:
BEVAN, John Innes Centre, Norwich, GB
PARTICIPANTS:
S.W.J. BRIGHT, ICI Seeds, Warefield, GB
G. COUPLAND, John Innes Centre, Norwich, GB
J. DANGL, Max-Planck-Institut, Köln, D
C. DEAN, John Innes Centre, Norwich, GB
M. DELSENY, CNRS, Univ. Perpignan, Perpignan, F
P. GALLOIS, CNRS, Univ. Perpigna, Perpignan F
J. GIRAUDAT, Inst. Sciences Végétales, CNRS, Gif-sur-Yvette, F
P.J.J. HOOYKAAS, Clusius Laboratories, Leiden, NL
C.M. KARSSEN, Agricultural Univ., Wageningen, NL
T. KAVANAGH, Trinity College, Dublin, IRL
C. KONCZ, Max-Planck-Institut, Köln, D
A. KOOL, Zaadunie, Enkhuizen, NL
M. KOORNEEF, Wageningen Agric. Univ., Wageningen, NL
J. LEEMANS, Plant Genetic Systems, Gent, B
K. LINDSEY, Univ. of Leicester, Leicester, GB
J.M. MARTÍNEZ-ZAPATER, CIT-INIA, Madrid, E
R. MASTERSON, Max-Planck-Institut, Köln, D
P. MEYER, Max-Planck-Institut, Köln, D
B. MULLIGAN, Univ. Nottingham, Nottingham, GB
P. OLESEN, Danisco A/S, Copenhagen, DK
M. PAGES, CSIC, Barcelona, E
J. PASZKOWSKI, Technische Hochschule, Zürich, CH
A. PEREIRA, Centre for Plant Breeding & Reprod. Res., Wageningen, NL
B. REISS, Max-Planck-Institut, Köln, D
P. VAN DEN ELZEN, Mogen, Leiden, NL
H. van MONTAGU, RU Gent, Gent, B
L. WILLMITZER, IGF Berlin, Berlin, D

BACKGROUND INFORMATION

The principal goal of the T project *Molecular identification of new plant genes* was to develop methods for identifying and isolating genes of agricultural importance, using the benefits of the model plant *Arabidopsis thaliana*, and apply these methods to understanding and manipulating areas of plant development and physiology important in plant productivity. In the process of achieving these goals, the T project also aimed to establish and maintain, over a period beyond the scope of the project, an infrastructure capable of sustaining future pan-european progress in *Arabidopsis* research. A subsidiary aim was to set up an industrial platform to whom the findings of the work would be made available and from whom scientists could gain ideas and support for applying their studies for the benefit of EU biotechnology and agriculture.

OBJECTIVES AND PRIMARY APPROACHES

The 'T' project provided a framework that linked scientists working on methods for gene identification to scientists dedicated to answering fundamental quetions in plant biology using genetics and physiology as their disciplines. This linkage aimed to accelerate the dissemination of techniques for gene seach, and their

applications to the molecular characterization of genes involved in flowering, seed development and embryogenesis.

A. PHYSICAL MAPPING

SUBCOORDINATOR:
C. DEAN, John Innes Centre, Norwich, GB

RESULTS AND DISCUSSION

1. Physical mapping of the *Arabidopsis* genome (C. DEAN, John Innes Centre, Norwich, GB)

New markers with which to screen 5 YAC libraries have been obtained. We estimate through using this increased number of markers and libraries that we have a genome coverage of chr 4 and 5 of 80%, assuming a genome size of 100Mb evenly split between the five chromosomes. In addition to maximising the density of RFLP markers used against the YAC libraries, we undertook chromosome walking experiments on a limited genomic region (15cM region of chromosome 4). this region is covered by three large YAC contigs (involving 110 clones) covering almost 3.5Mb. Analysis of these contigs revealed that at least 20% of the YAC clones involved are chimaeric. We have also identified YAC clones carrying some of the major classes of repeated DNA sequences, such as ribosomal DNA, chloroplast DNA and the centromeric heterochromatin repeat (termed *Hind*III repeat). YAC clones carrying the *Hind*III repeat exhibited a high degree of instability.

2. Construction of an *Arabidopsis* genomic library in Phage P1 (T. KAVANAGH, Trinity College, Dublin, IRL)

We are currently attempting to generate a Phage P1 library of *Arabidopsis* genomic DNA. The task involves:

(i) preparation of megabase genomic DNA,
(ii) partial digestion of megabase DNA to a size range of 90-100kb,
(iii) ligation of size-fractioned DNA with Phage P1 vector,
(iv) packaging and transfection of ligation products,
(v) characterization of clones.

We are producing megabase DNA by embedding *Arabidopsis* protoplast preparations in agarose followed by digestion with Proteinase K in the presence of SDS and EDTA. DNA prepared in this way has a molecular weight of several megabases. We have used partial digestion of embedded DNA with SAU3A at 4°C overnight to generate DNA fragments in the size range 85-120kb. A major problem however is obtaining sufficient quantities of DNA in this size range for ligation with BamH I-digested Phage P1 vector DNA (pAd10-SacBII). We prepare our own packaging extracts: these gave packaging efficiencies 60 fold higher than the commercial extracts with concatamerized vector.

3. Construction of an integrated genetic and physical map at the *Arabidopsis* genome (P. VOS & M. ZABEAU, Keygene, NL)

The aim of this project was to develop a detailed integrated genetic and physical map of the *Arabidopsis* genome. AFLP methods provided at least a 1000 informative amplified fragments between Colombia and Ler, and 150 markers have been mapped onto Recominant Inbred lines. At least 500 additional markers will be mapped in the future. A YAC library prepared by CEPH which contains large

inserts will be screened to position YAC clones on the genome. Presently AFLP markers are being used to assess the library.

B. RESOURCE CENTRES

SUBCOORDINATOR:
B. MULLIGAN, University of Nottingham, Nottingham, GB

1. The *Arabidopsis* Resource Centre (B. MULLIGAN / M. ANDERSON, Univ. of Nottingham, Nottingham, GB)

Approximately 8,000 *Arabidopsis* accessions are now held in Nottingham. All stocks are well used with recombinant inbred lines and T-DNA transformed pools being particularly in demand. At present, approximately 1,000 tubes of seed are shipped per month. An acquisition of T-DNA transformed lines from K. Lindsey (Leicester) is currently being bulked for distribution, as is the Redei collection of mutants, received from the Ohio Stock Centre. Information about stocks is disseminated through the Seed List, of which more than 500 copies are mailed annually, the Electronic Seed List, accessible through the *Arabidopsis* Research Companion Gopher server and through AAtDB. All stock information is now incorporated at regular intervals into the *Arabidopsis* database AAtDB and AIMS. We have found that the best method of electronic storage of images of stocks held is on CD. An in-house database covering all aspects of acquisition history, quality control etc. has now been established to streamline Stock Centre administration.

2. *Arabidopsis* DNA Resource Centre (J. DANGL, Max-Planck-Institut, Köln, D)

We developed a useful routine for storage and distribution of libraries and probes around the world (around 2 days from request to shipping; shipping via courier). I have participated in numerous discussions aimed at positioning a DNA Resource Centre component as a fundamental part of the EEC's long term goals in *Arabidopsis* research. Two complementary YAC libraries, those most commonly used by the international *Arabidopsis* community, are also stocked and available from the DNA Resource Centre. A third library, having the advantage of a different method of preparation, but the disadvantage of smaller average insert size is also available if needed. Dr. Caroline Dean's laboratory provides a back-up storage facility.

In current reporting period (1993), we have filled the following request of our materials:

Phage RFLP markers:	7 times
Cosmid RFLP marker:	10 times
ARMS RFLP probe set:	5 times (A NEW ITEM; thanks to Dr. Toni Schaffner, München)
Genomic libraries:	22 times
cDNA libraries	30 times
YAC libraries:	3 times, 2 times selected clones
TOTAL:	**77 requests.**

C. GENE REPLACEMENT

SUBCOORDINATOR:
P. VAN DEN ELZEN, Leiden, NL

1. Site-directed mutagenesis of endogenous genes in *Arabidopsis thaliana* (P. VAN DEN ELZEN, Mogen, Leiden, NL)

The strategy involving selecting potential recombinants on the drug alpha amanitin was not feasible due to the insensitivity of *Arabidopsis* cells to this drug. We decided to continue to target the rpII215 gene but with an alternative strategy using a very large amount of homology in the targeting vector and relies on a robust PCR screen for recombinants. This method will be more time consuming than a selection system but offers the advantage this it is applicable to any plant gene. The flanking regions of the rpII215 clones have been subcloned and a restriction map of these has been determined for 11 different restriction enzymes. This data has been used to design a cloning strategy for creating a targeting vector which provides 10.5kb of homology on one arm of the vector and 7.8kb of homology on the other arm. An nptII cassette which does not contain any *SmaI* sites has been made for insertion into the above clones. The construction of the targeting vector will be completed with the joining of two arms and the subsequent cloning of the construct into a binary vector for transformation.

2. Site-directed mutagenesis in *Arabidopsis thaliana* using *Agrobacterium* as a gene delivery system (P.J.J. HOOYKAAS, Clusius Laboratories, Leiden, NL)

The objective of this project is to develop a highly efficient transformation system for gene targeting using *Agrobacterium* as DNA donor and establish variables and frequencies of gene repair.

Much progress was made in transformation of *Arabidopsis* root explants via co-cultivation with *Agrobacterium*. Different parameters, such as plant growth conditions (light, temperature), induction with acetosyringone, the use of hypervirulent *Agrobacterium* strains as well as tissue culture conditions, were studied to optimize transformation of ecotype C24. By using a construct with a gusA reporter gene and both *neomycin-* and *hygromycin phosphotransferase (nptII and hpt resp.)* genes as plant selectable markers the transformation frequency was optimized resulting in appr. 4.5 GUS positive cells per root explant (= 3-5 mm). This number correlates to stable antibiotic resistance of 2 to 3 calli per root explant. The regeneration frequency of resistant calli reaches 60 to 80%.

Homologous recombination experiments using primary hygromycin resistant transformants containing a defective nptII-gusA fusion gene are being analysed. Due to problems with hygromycin selection it was difficult to obtain these primary transformants, but now gene targeting experiments will be started soon. The frequency of recombination will be studied varying parameters such as length of homology and vector type (integration- versus replacement-type).

3. Gene targeting via YACs (P. MEYER, Max-Planck-Institut, Köln, D)

The goal of this project is the mutation of a transgene by homologous recombination, using a YAC with a sufficiently large homology that might increase the usually low levels of homologous recombination. To test for selection markers we have generated an *Arabidopsis* transformant that contains one copy of a plasmid carrying the TK-gene from Herpes simplex and the tms 2 gene from *Agrobacterium tumefaciens*. We found that, in contrast to animal cells, it is not feasible to counterselect TK activity by Gangcyclovir. Counterselection of tms 2 activity by NAM was

possible when seedlings or calli from the transformant were cultured om 3μM NAM. Reconstitution experiments revealed that a selection for the loss of tms 2 activity in protoplast was not sufficient to select very rare events. To isolate a suitable YAC clone for recombination cloning of 300-400 kb NotI-fragments into pYAC45 was initiated. YAC containing a deleted tms 2 gene will be generated by homologous recombination in yeast. The transfer of YACs into protoplasts has been tested with a yeast strain carrying a pYAC45 vector containing a GUS gene with an intron that is only spliced out in the plant. We succeeded in fusing tobacco protoplast with nuclei of the recombinant yeast strain which gave a significant GUS expression, probably due to the transfer, transcription and correct splicing of the GUS gene. At present a YAC of 100 kb is prepared that contains a GUS gene and a Hyg gene, to test whether larger YACs are also efficiently transferred by this technique and whether resistant clones can be selected after integration of the Hyg gene.

4. Gene Replacement (L. WILLMITZER, IGF Berlin, Berlin, D)

A model system for gene targeting studies consisting of a non-functional *hpt* gene under control of the CaMV 35S promotor and provided with the *ocs* polyadenylation signal which was stably inserted into the *Arabidopsis* genome as chromosomal target for gene replacement by the intact *hpt* coding region was devised. This targeting DNA was later on introduced by PEG-medicated direct DNA uptake into protoplasts. A targeting frequency of about 3×10^{-4} was determined by selection for hygromycin resistant transformants that were further analysed by PCR and by Southern blot experiments. In order to test the influence of enhanced length of homology between target and targeting DNA, several attempts were made to clone *hpt* genes together with the largest possible region of flanking plant DNA from a locus for which there was a precedent for successful gene targeting. However, due to instability of the lambda and cosmid clones isolated from genomic libraries prepared from both a parental line, only strongly rearranged DNA fragments were obtained making them unusable for further gene targeting experiments. Due to the lack of suitable negative selectable markers for plant protoplasts, different *hpt* antisense genes in conjunction with sense *hpt* were tested for negative selection. When stably introduced into protoplasts by PEG-mediated direct DNA uptake, these constructs showed two to three fold reduced relative transformation frequencies as compared to control constructs that do not promote *hpt* antisense expression. The observed reduction, however, is not sufficient for a negative selection marker appropriate to strongly enrich for gene targeting events. In order to further follow the strategy of offering long streches of homologous DNA so as to promote homologous recombination, we decided to test an entirely new system for cloning and stable maintaining large DNA molecules in a bacterial cloning system, bacterial artificial chromosomes (BAC), for its applicability to clone large fragments of *Arabidopsis* DNA.

5. Stimulation of recombination by recA (B. REISS, Max-Planck-Inst., Köln, D)

Plants transgenic for an *E. coli recA* gene fusion to the SV40-T nuclear targeting sequence expressed recA protein exclusively and in high amounts in the nucleus. Using biochemical assays, this protein has been shown previously to be active in all aspects relevant for promoting recombination. Several types of experiments were set up to investigate the activity of recA in these plants using plants from other participants. Tests on extrachromosomal recombination, gene targeting frequencies, and extrachromosomal recombination demonstrated that recA in these

plants is fully functional, able to interact with the plant recombination machinery, and stimulates recombination.

D. GENE TAGGING

SUBCOORDINATOR:
G. COUPLAND, John Innes Centre, Norwich, GB

1. Establishing a transposon-tagging system in *Arabidopsis* by use of the maize transposon *Ac* (G. COUPLAND, John Innes Centre, Norwich, GB)

We have now identified 28 heritable mutant phenotypes in the transposon populations derived from Ac and Ds crosses as described in previous reports. Twenty of these were tested to determine whether the mutation and the transposed *Ds* element are genetically linked. If the mutation is within 5 cM of the *Ds* element, it is referred to as linked. In 11 cases the mutation and the transposed *Ds* element were not linked, and therefore these mutations were not tagged. The remaining 9 mutations that were tested were linked to the *Ds* element, however in two of these more detailed mapping experiments identified cross-overs separating the *Ds* and the mutation. Many of the seven mutations that were inseparable from *Ds* are being worked on in detail. Two of them are certainly tagged *(albino3, curly leaf)* and the genes have been cloned. For a further 4 (two dwarfs, an early flowering and a late flowering) the mapping populations were greatly extended, revertants are being sought and DNA fragments adjacent to the *Ds* element in each line were cloned. The usefulness of the strategy is illustrated by how interesting and unexpected the two tagged genes are. The *curly leaf* gene is required to inhibit the expression of the flower homeotic gene *Agamous* both in the leaves and flowers, a function which was not previously expected to be required in the leaves. Moreover, a 100 amino acid region within the *Albino3* protein shows 40% identity to an unusual bacterial protein and therefore indicates novel functions required for chloroplast biogenesis.

2. T-DNA insertional mutagenesis (C. KONCZ, Max-Planck-Institut, Köln, D)

A current problem of T-DNA insertional mutagenesis is that a considerable proportion of T-DNA transformed plants segregate mutations displaying with T-DNA inserts. Therefore, a detailed genetic and molecular characterisation of mutant collection was initiatied, to assess the efficiency of the T-DNA insertional mutagenesis technique. Enrichment, screening and genetic analysis of 1,400 T-DNA insertional lines was completed in 1992, and a part of this collection was characterised by the *Arabidopsis Stock Centre*, and made available for public use. 20 mutants identified by screening in soil and Petri dishes using various conditions were subjected to detailed genetic linkage and molecular analyses. In 3 of these mutants, the T-DNA was still found in position, allowing an immediate characterisation of corresponding genes. In other mutants, the mutations showed either an independent segregation from the T-DNA(s), or were found in tight linkage with a T-DNA insert located on the same chromosome. In several instances T-DNA inserts displayed meiotic instability when outcrossed with wild type. To study the mechanism of T-DNA instability, complex T-DNA inserts, carrying 2 or 3 tandem T-DNA copies in the same locus were re-isolated from mutants that showed allelism with the *erecta and pinformed loci*, respectively. A gradual deletion of T-DNA inserts and flanking plant DNA sequences was detected in F2 progeny obtained by recurrent outcrosses of these mutants. Although the mechanism generating these deletions is still unclear, the available data suggest that

intrachromosomal recombination, and possibly alteration of the methylation of plant DNA's flanking the T-DNA inserts are major factors that affect the stability of complex T-DNA insertions in *Arabidopsis*. One of the insertional mutants, which displayed wide pleiotropic effects on growth, was shown to be in a novel β-transducin gene.

3. **Characterisation of an *Activator/Dissociation (Ac/Ds)* transposable element system in *Arabidopsis thaliana*** (R. MASTERSON, Max-Planck-Institut, Köln, D)

82 transgenic Arabidopsis containing Ac,DsDHRF and Ds35S elements, introduced by *Agrobacterium*-mediated transformation, have been analysed to identify plants containing single and multiple T-DNA copies. Current efforts have been directed towards screening large Ac/Ds populations in which a novel 'mobile promoter' type element, termed Ds35S, has been used to tag gain of function mutations. Over-expression has occurred in some Ac/Ds35S2 plants, which may result from the activation of host genes by the CaMV35S promoter located in the Ds35S2 element.

4. **Transposon tagging in *Arabidopsis thaliana*** (M. van MONTAGU, RU Gent, Gent, B)

Activation of Ds upon transformation with an *Ac* construct, containing the transposage cloned behind the Tr1' promoter, was measured by the formation of green calli on kanamycin-containing medium the range of Ds excision varied from 5-20% to 40-80%. Lines with high excision frequency will be used further for mutation induction. The Selected Restriction Fragment Amplification (SRFA) technology is being used to facilitate the cloning of genes based on locus map position. As a test case for *Arabidopsis*, SRFA fragments are being determined closely linked with the TRN (TORNADO) locus. They will allow to land close by the gene thus avoiding chromosome walking. AdTph-specific primer and a restriction site primer for amplification result in fragments that harbour part of the transposable element. We expect that it will be relatively easy to identify fragments co-segregating with the transposon-induced mutation.

5. **Enhancer detection with the En-I transposable element system** (A. PEREIRA, Centre for Plant Breeding and Reproduction Research, Wageningen, NL)

We demonstrated the effectiveness of the En-I system by tagging and cloning the MS2 gene from *Arabidopsis*. We found that the transposition frequency is constant over several generations and that the frequency of new transpositions in a progeny is about 10 to 15% of all *I* elements inserts.
Mapped *I* elements are more or less randomly distributed over the chromosomes, although we have found some clustered inserts. The unmapped transposons will be placed on the physical (YAC) map. Lines containing mapped *I* elements will soon be distributed to the Nottingham *Arabidopsis* Stock Centre. Apart from the ms2 mutant we have found several putants, such as a waxless (*eceriferum*) mutant, which segregated with the homozygous presence of an *I* element and in the presence of *En-transposase*, somatic reversion to wildtype seems to take place. Cloning flanking DNA from the candidate *I* element insert is in progress. For detecting enhancer sequences with an *I* element containing a minimal promoter-GUS fusion, we made several crosses between enhancer-detector plants and *En-transposase* plants. So far we did not detect mobility of *I* elements in the F1 generation.

6. Gene Tagging (L. WILLMITZER, IGF Berlin GmbH, Berlin, D

Our main objective was the establishment of various selection procedures to identify and enrich for plants in which wild-type Ac had excised from its original position and integrated elsewhere in the genome. Also a two component Ac/Ds system was developed that consisted of an immobilized Ac element and a non-autonomous, Ds-like element, DsA. Both systems allowed the selection of plants with excision/transposition events of the mobile element fixed in the germline. A screening program for transposon induced mutants showing visible mutant phenotypes was carried out on progenies of 43 Ac families and 559 Ac^{c12}/DsA families. Mutants were indentified in 48 (all derived from Ac^{c12}/DsA families) of the 602 families analyzed, and showed a wide variety of phenotypes. Four putative DsA insertion mutants included two with embryo lethality, one with dark green unexpanded rosette leaves, and one with curly growth of stems, leaves and siliques. The new DsA insertions will be stabilised to allow a more rigorous cosegregation analysis and to facilitate future efforts to reclone the transposed DsA with flanking plant DNA. Cosegregration analysis performed on an embryo lethal mutant revealed 5 out of 45 plants recombinant, indicating DsA insertion near to but not into the mutant gene responsible for the embryo lethal phenotype.

E. FLORAL INDUCTION

SUBCOORDINATOR:
J. MARTINEZ-ZAPATER

1. Identification and characterisation of genes involved in floral induction (J.M. MARTINEZ-ZAPATER, CIT-INIA, Madrid, ES)

We have constructed a YAC contig that includes the chromosal region containing the FVE locus, which regulates floral initiation in response to daylength, and a cosmid contig overlapping the smallest YAC covering the gene is being made. Three genes segregating for late flowering in the crosses have caused multiple difficulties in the detection of real combinants in the region surrounding *FVE* since most of the plants with a recombinant phenotype did not have the expected genotype at the *FVE* locus. To further confirm the chromosomal location of *FVE* we are currently mapping it with respect to a Ds Insertion linked to RFLP marker 1251 on chromosome 2. Deletion mutagenesis has been carried out, and we have screened 55,000 M2 plants and isolated more than 40 later flowering mutants. These mutants are being ordered in several phenotypic groups depending on the effect of different environmental variable on the late flowering phenotypes, and those that show a similar phenotype as *FVE* mutants are tested by complementation analysis with the two EMS alleles already identified at this locus. Moreover, the Ds insertion on chromosome 2 is being used to identified Ds tagged late flowering mutants at the *FVE* locus.

2. Isolation of the *CO* locus and identification of other mutations affecting flowering time in *Arabidopsis* (G. COUPLAND, John Innes Centre, Norwich, GB)

The *co* mutation defines a gene that is required for early flowering under long days. The gene has been located on a 6kb fragment in a 1700kb YAC contig on the top arm of chromosome 5. The major effort over the last year was to identify the gene within this region. Two approaches were used: screening for cDNAs encoded by the region and sequencing of the genomic DNA to try to identify candidate open reading frames (ORFs). One cDNA was isolated and sequenced, but only approximately half of this is encoded by the 6kb that complements the

co mutation. The entire 6kb fragment that complements the mutation was therefore sequenced, and surprisingly a second gene that is highly homologous to the cDNA was identified. This second gene is the only one that is completely contained within the complementing region. The polymerase chain reaction was used to amplify this gene from 3 *co* mutants. Two independent amplifications were performed on each mutant to exclude the possibility that any of the errors were caused during the PCR. The genes derived from all three mutants were shown to contain mutations: two EMS-induced alleles contained base substitutions, and the original X-ray allele was a 9 bp deletion.

3. Floral induction in *Arabidopsis* (C. DEAN, John Innes Centre, Norwich, GB)

RFLP experiments enabled us to map the FCA locus between the two RFLP markers LEM 580 and 226. These RFLP markers were hybridized to the different YAC libraries to build contigs on either side of the FCA locus. Analysis of recombinants with closely linked probes and restriction maps of this genomic region, currently places the FCA locus in a 30kb region. One of the *fca* alleles, induced by fast-neutron mutagenesis has a large rearrangement in this genomic region, with one end of that arrangement ending in a 1.9kb BamHI fragment. This gives us evidence that FCA gene lies on this fragment. Complementation experiments have provisionally shown the location of the FCA gene and cDNA identification is now in progress.

4. Floral Induction in *Arabidopsis thaliana* (M. KOORNNEEF, Wageningen Agricultural University, NL)

Double mutants have been obtained and confirmed by allelism tests to be double mutants for 35 of the 45 possible combinations of the 10 loci under investigation. Physiological analysis will be initiated in the coming months. The *hy2* and *hy3* and the *hy2,hy3* double mutant has been combined with the late flowering mutants *co, fra, gi* and *fwa*. In these double mutants the earlyness of the *hy* mutants, affected in the phytochrome system, resulted in earlyness in the late flowering background, although these double mutants are more late than the *hy* mutants in wild type background. The flowering behaviour of all these genotypes in light conditions differing in their Red/Far-red ratio indicate that the effect of light quality is not only due to phytochrome B. When flowering of these double mutants is compared in long days and short days the day length sensitivity is not abolished as it almost is in the *hy* mutants in wild type background. This suggests that the day length effect might not be meditated by phytochrome but that phytochrome as such plays an inhibitory role in flowering. The analysis of the flowering time genes not allelic to the previously described loci has been completed. It appears that the extremely late *F* mutant and the mutants isolated by the group of Cetl and Relichova are all allelic with the *FLA* = *FRI* locus identified in very late ecotypes and located on the top of chromosome 4. It was found that the dominant allele confers extreme lateness only in combination with dominant alleles at the *FLC* locus located just above *tt4* on the top of Chromosome 5. It was found that all ecotypes tested (Col, Li, Est, S96 and Ws) have dominant alleles at this locus with the exception of Ler. The *FIC* locus or a closely linked gene present in the same ecotypes gives lateness in combination with recessive alleles at the *id* locus identified by Redei. Mutants at the *fwa* locus are dominant and are relatively insensitive to vernalization but responsive to day length. The gene has been identified on a cosmid and lambda clones provided by Jeff Leung. A number of cDNAs corresponding with this cosmid has been isolated. Since we found that increasing the number of wild types copies using trisomics for chromosome 4, resulted in earliness, the (semi-domi-

nant) *fwa* mutant is probably not a gain of function mutation and therefore attempts to complement the *fwa* mutant with the subcloned cosmid are underway.

5. Molecular characterisation of the *APETALATA 2* gene (M. van MONTAGU, RU Gent, Gent, B)

In situ hybridisation studies indicate that AP2 mRNA accumulates early in floral bud initiation and this is independent of the action of LFY and AP1 genes. The AP3 and PI loci positively regulate AP2 expression late in flower development, and AG and AP2 expression patterns overlap. Sequence data has been obtained for 5 AP2 alleles. Defects in seed coat formation caused by AP2 are being investigated in crosses between AP2, GL2 and ttg mutants. The seeds of these progeny are being examined microscopically. In future the role of AP2 in late embryogenesis will be examined further.

F. SEED DEVELOPMENT

SUBCOORDINATOR:
M. KOORNNEEF

1. Acquisition of desiccation tolerance in developing seeds of *Arabidopsis thaliana* (C.M. KARSSEN, Agricultural University, Wageningen, NL)

It was shown that desiccation tolerance can be induced in double mutant seeds by either *in vivo* treatment with the ABA analog LAB 173 711 or *in vitro* by treatment of unripe seeds with sugars and/or ABA. In addition, we recently showed that desiccation tolerance can be induced by slow and well controlled drying of immature double mutant seeds. This ABA-insensitive double mutant, in which no ABA action is detected, offers the unique possibility to study acquisition of desiccation tolerance without interference of any residual endogenous ABA. Recent experiments suggest that besides an ABA-induced desiccation tolerance, another ABA-independant pathway for induction of desiccation tolerance exists. Detailed Northern and Western analysis of double mutant seeds in which desiccation tolerance has been induced by the different above mentioned methods is currently under investigation.

2. Seed development and Abscisic Acid (M. KOORNNEEF, Wageningen Agricultural University, Wageningen, NL)

The physiological and genetic characterisation of two non-allelic mutants with reduced seed dormancy has been completed. These mutants do not show clear pleiotropic effects and are as sensitive as the wild type to abscisic acid (ABA). The mutants show an additive effect with *aba* and *abi3* mutants for dormancy and seed longevity indicating that genes specific for dormancy do exist. A mutant with 'green' seed phenotype somewhat similar to that of extreme *abi3* alleles but not allelic has been isolated. In this mutant seed longevity is normal and 2S storage protein accumulation is not reduced. Screens for mutations affecting some aspect of seed dormancy has been applied to the progenies of 5,000 transgenic plants from the Feldmann collection. 10 mutants have been isolated and co-segregation of the dormancy trait and kanamycin resistance is underway and the cloning of genes in tagged mutants will be initiated. Analysis of the 'wilting' mutants resistant to the gibberellin biosynthesis inhibitor paclobutrazol indicated that two of the mutants were ABA insensitive. However, some mutants are putative candidates for new aba loci since these mutants are not ABA insensitive. Preliminary experiments indicate that the Ler ecotype is more dormant that the Col ecotype. The genetics

of this difference will be investigated using the recombinant inbred lines derived from the cross of these two ecotypes by Lister and Dean. EM2 promotor-GUS fusions have been combined with various *aba* and *abi* mutants and will be used to investigate the effect of these mutants on this promotor. *Arabidopsis* containing a RAB17-GUS fusion was mutagenized and 5000 M1 plants were individually harvested. This mutagenized population will be used to isolate mutants that enhance or suppress the expression of this promotor. The maternally inherited heart shaped seed mutant *ats* (aberrant seed shape), mapped to the middle of chromosome 5 and has been characterised, and it was shown that the aberrant shape is established in the first few days after fertilisation and is due to the lack of two cell-layers of the testa. The resulting three cell layers show characteristics of both the inner and outer integuments. The mutant has a reduced seed coat imposed dormancy. This mutant provides the first experimental evidence of the role of the seed coat in the determination of the seed shape. Screening M2 seeds resulted in the isolation of two mutants with larger seeds and a detailed analysis of large seed genotypes is in progress.

3. Expression of ABA-responsive genes in *Arabidopsis* (M. PAGES, CSIC, Barcelona, ES)

Expression of the maize rab17-GUS gene in *Arabidopsis* wild type and ABA-deficient and ABA-insensitive mutants shows that alternative pathways are responsible for seed specific expression and ABA-osmotic stress induction of the rab promoter. Rab 17/GUS-expressing *Arabidopsis* have been mutagenized and are under screening to select plants with an altered expression of the reported gene.

Expression and subcellular localisation of the RNA-binding protein from maize (MA16) has been studied in transgenic *Arabidopsis* plants carrying the 35S/MA16 construct. We studied the *in situ* hybridisation pattern of the Em-like *Arabidopsis* gene, and we are sequencing and characterisation the *Arabidopsis* homolog of the maize rab28 gene. We have also sequenced one genomic clone showing strong hybridisation with the rlea2 cDNA.

4. ABA-responsive genes expressed during late seed development (M. DELSENY, CNRS, Univ. Perpignan, Perpignan, F)

We demonstrated that a 249 bp sequence of the AtEm1 promoter was sufficient to confer ABA responsiveness and that in the upstream region there was an enhancer element. Expression time course of Em1-GUS gene mimics the expression of the endogenous gene which is specific for the late embryo. We further demonstrated the dependence of AtEm1 upon a functional AB13 protein for expression. Meanwhile we have initiated long range sequencing across the AtEm1 locus, which was located on chromosome #3. We isolated two related cDNA, AtM10 and AtM17 which cross-hybridized with an incomplete radish cDNA, AtM10 codes for a short protein with a single C-terminal domain rich in cysteines reparated by glycine-rich spacer. AtM17 is cDNA with 4 repeats of the cysteine motif, a situation similar to the radish cDNA. We also made a dry seed cDNA library from which we randomly sequenced slightly more than 100 clones. Using this strategy we could identify most of the so called LEA cDNA and particularly homologues to RAB17 and RAB28.

5. Analysis of the *abi3* and *abi1* genes and the mode of action of genes involved in drought resistance (J. GIRAUDAT, Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, F)

The expression pattern of the ABI3 gene has been established using both Northern analysis and transgenic *Arabidopsis* plants carrying ABI3-GUS transcriptional fusions. We have also analyzed the ability of the AB13 gene product to mediate abscisic acid (ABA)-induced gene expression by characterising *Arabidopsis* transgenic plants ectopically expressing the ABI3 gene in vegetative tissues. Finally, we have investigated further the respective roles of ABA and of the ABI gene product in regulating gene expression during silique development. We have then analysed the effects of the severe abi3-4 mutation and that of the aba (ABA biosynthetic mutant) mutation on these various programs of gene expression. We have now reached the final stages of our positional cloning of the abi1 gene; RFLP analysis has allowed us to identify a 150 kb region containing the abi1 locus. We have identified a 6kb fragment containing the functional abi1 gene. Final identification of the target gene involves physical characterisation of the corresponding genomic and/or cDNA clones in various available abi1 alleles. The nucleotide sequence of cDNA clones that either accumulate (3) or decrease (1) in root systems of *Arabidopsis* plants subjected to progressive drought, has been characterised. We have also analysed the kinetics of expression of the corresponding mRNAs in both roots and leaves of wild-type plants during progressive drought. The responsiveness of these genes to exogenous ABA has been determined in seedlings of wild type and abi mutants. Finally, the involvement of endogenous ABA in regulating the abundance of these transcripts during progressive drought has been determined by comparing their expression levels in wild type, abi and aba mutants.

G. EMBRYOGENESIS

SUBCOORDINATOR:
K. LINDSEY

1. Insertional Mutagenesis to Isolate Genes Controlling Embryogenesis in *Arabidopsis* (K. LINDSEY University of Leicester, Leicester, GB)

Over 2000 independent transformants of *Arabidopsis* containing PΔgusBin19, a promoter trap vector designed to tag genes by the random generation of gene fusions *in vivo*, have been made. Three lines that exhibit GUS fusion activity in seeds have been identified for more detailed analysis: these are designated AtEM-101, AtEM-201 and AtEN-101. Embryonic pattern mutants have also been identified. In AtEM-101 GUS activity is restricted to the basal region of the embryo and root tip; in AtEM-201 GUS activity during the heart-stage is found throughout the entire embryo. In seedlings activity is found in root tips and also at lower levels in cotyledons and hypocotyl. T-DNA 5' flanking sequences have been amplified by IPCR, and cDNAs have been isolated and are being analysed further. Expression in line AtEN-101 was restricted to the endosperm, and transiently in the globular embryo. GUS activity is found in the root cap and elongation and maturation zones, but not the meristem. A number of lines that carry mutations affecting embryogenesis or seed development have been identified, which fall into three classes: *embryonic lethals*, which fall to complete embryogenesis; *altered pigmentation*, including albinos, *fusca* and *transparent testa* line; and *seedling defectives*, which are affected in tissue pattern formation.

2. Tagging genes involved in embryogenesis (P. GALLOIS, CNRS URA 565, Univ. Perpignan, Perpignan, F)

Molecular studies have been initiated with T-DNA line 276-1 which has been previously shown to express GUS in the suspensor only and to contain a single copy of the GUS gene in its genome. A 1480 bp fragment upstream of the GUS ATG initiation codon has been isolated and is present as a single copy in the wild type genome. Northern analysis was carried out on total RNA extracted from different organs showed a 1500-1600 bp transcript only in RNA from the tissues where the suspensor is present. This indicated that the IPCR fragment corresponds at least partially to a transcribed sequence. cDNA clones have been isolated using the upstream fragment are being sequenced and characterised and *in situ* studies will be conducted on the suspensor cells.

3. Detection of GUS-mRNA in Transgenic *Arabidopsis* Embryos (P. OLESEN, Danisco A/S, Copenhagen, DK)

A ^{35}S-labelled riboprobe set (sense and antisense) (approx. 250 bp) was constructed against a specific region of the GUS-gene. This worked with a minor adjustment of the ISH-method of J. Langdale (Oxford). ISH-experiments were conducted with the GUS-probes on GUS-positive *Arabidopsis* embryos, and results were obtained with 3 'promoterless' GUS-transgenic lines from Leicester as well as transgenic *Arabidopsis* lines with 35S promoter in front of GUS-gene. A weak background was found with the sense control probe and in the GUS-negative control embryos. The GUS-enzymatic expressions of the 'promoterless' lines (M-1, M-2 and M-3) were confirmed and extended to more embryo developmental stages by these ISH-experiments. A root and suspensor specific GUS-positive line from Perpignan was investigated without achieving clear positive results. The GUS enzymatic expressions of the suspensorline were verified and possibly extended to weak expressions in the basic part of the embryo as well as in the endosperm. A ^{35}S-labelled riboprobe set (sense and antisense) (approx. 500 bp) was constructed against the M-1 gene from Leicester. Probes against the M-1 gene gave positive ISH-results in both sections of embryos from the M-1 *Arabidopsis* line as well as in sections of embryos from wild type *Arabidopsis*. Thus we were able to confirm that the GUS-gene was insertet into or near by an embryo specific *Arabidopsis* gene in the M-1 line from Leicester.

MAJOR SCIENTIFIC BREAKTHROUGHS

At the end of the first two and a half years of the project, many of the goals described above have come to fruition, particularly those involved in methods development, such as transposon tagging and map-based gene cloning, and in establishing resource centres. Significant progress has also been made in incorporating molecular genetics techniques into established areas of plant physiology, such as flowering and seed development, and a promising start has been made in the systematic isolation and characterisation of genes involved in these important processes. The EC-sponsored work on physical mapping has been central to a world-wide effort to establish a map of the entire genome, and the success of this work has led several participants in the BRIDGE programme to instigate, in the framework of the BIOTECH programme, the first attempts to sequence the entire *Arabidopsis* genome. The EC projects in *Arabidopsis* genetics have a clearly defined mission and a high international profile, reflecting the success of the strategy of integrating national *Arabidopsis* programmes throughout Europe. The teams set up in BRIDGE have become a productive and forward looking group, capable of extending their work into diverse aspects of agricultural and biotechnological

research in the EU well beyong the original remit of *Arabidopsis* genetics. This is the greatest and most enduring achievement of the scientists involved in the BRIDGE programme.

MAJOR COOPERATIVE LINKS

The most important links developed have been those involved in coordinating the European *Arabidopsis* Stock Centre at Nottingham and the Ohio Stock Centre in the US. The EU contribution to the multinational effort in physical mapping has been pivotal. Finally, the establishment of the Plant Industrial Platform will provide an effective way of transferring the results of *Arabidopsis* research to the agricultural and biotechnological sectors of EU industry.

PUBLICATIONS

Joint publications

M. Devic, M. Delseny, P. Gallois, V. Hecht and K. Lindsey. Promoter trapping as a tool to study zygotic embryogenesis in *Arabidopsis thaliana* (submitted).

P. Gaubier, M. Raynal, G. Hull, G.M. Huestis, F. Grellet, C. Arenas, M. Pages and M. Delseny (1993). Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. Mol. Gen. Genet. **238**: 409-418.

P. Gaubier, M. Raynal, G.M. Huestis, G. Hull, F. Grellet, C. Arenas, M. Pages and M. Delseny (1993). Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. Mol. Gen. Genet. **238**: 409-418.

P. Gaubier, G. Hull, M. Raynal, G. Huestis, F. Grellet, C. Arenas, M. Pages and M. Delseny (1993). Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. Mol. Gen. Genet. (in press).

J. Martinez-Zapater, G. Coupland, C. Dean and M. Koornneef (1993). The transition to flowering in *Arabidopsis*. In: *Arabidopsis*. E.M. Meyerowitz and C.R. Somerville (eds). Cold Spring Harbor Laboratory, Cold Spring Harbor (in press).

J.J.J. Ooms, K.M. Leon-Kloosterzeil, D. Bartels, M. Koorneef and C.M. Karssen (1993). Acquisition of desiccation tolerance and longevity in seeds of *Arabidopsis thaliana*. A comparative study using ABA-insensitive *abi3* mutants. Plant Physiol. **102**: 1185-1191.

J.J.J. Ooms, K.M. Léon-Kloosterziel, D. Bartels, M. Koornneef and C.M. Karssen. Acquisition of desiccation tolerance and seed longevity in seeds of *Arabidopsis thaliana*. A comparative study using ABA-insensitive *abi3* mutants. Plant Physiol. (in press).

J. Vilardell, J.M. Martinez-Zapater, A. Goday, C. Arenas, M. Pages (1991) .Regulation of the rab17 gene promoter in transgenic *Arabidopsis* wild-type, ABA-deficient and ABA-insensitive mutants. Plant. Mol. Biol. (in press).

Individual publications

J. Putterill, F. Robson, K. Lee and G. Coupland (1993). Chromosome walking with YAC clones in *Arabidopsis*: isolation of 1700 kb of contiguous DNA on chromosome 5, including a 300 kb region containing the flowering-time gene *CO* Mol. Gen. Genet. **239**: 145-157.

D. Long, M. Martin, E.Sundberg, J. Swinburne, P. Puangsomlee and G. Coupland (1993). The maize transposable element system *Ac/Ds* as a mutagen in *Arabidopsis*: Identification of an *albino* mutation induced by *Ds* insertion. Proc. Natl. Acad. Sci. USA **90**: 10370-10374.

D. Long, J. Swinburne, M. Martin, K. Wilson, E. Sundberg, K. Lee and G. Coupland (1993). Analysis of the frequency of inheritance of transposed *Ds* elements in *Arabidopsis* after activation by a CaMV 35S promoter fusion to the *Ac* transposase gene. Mol. Gen. Genet. **241**: 627-636.

C. Koncz, K. Nemeth, G.P. Redei and J. Schell (1992). T-DNA insertional mutagenesis. Plant. Mol. Biol. **20**: 963-976.

J. Schell, C. Koncz, A. Spena, R. Walden, K. Palme (1992). Characterisation of genes affecting plant growth and development. In Harnessing Biotechnology for the 21st Century, M.R. Ladisch and A. Bose (eds), American Chemical Society, Washington, pp. 496-499.

J. Schell, C. Koncz, A. Spena, K. Palme and R. Walden (1993). Genes involved in the control of growth and differentiation in plants. Gene **135**: 245-249.

C. Koncz, K. Nemeth, G.P. Redei, J. Schell (1993). Homology recognition during T-DNA integration into the plant genome. In: Homologous Recombination in Plants, J. Paszkowski (ed.), Kluwer Academic, Dordrecht (in press).
F. Jasper, C. Koncz, J. Schell and H.-H. Steinbiss (1993). *Agrobacterium* T-strand production *in vitro*: sequence specific cleavage and 5-protection of single-stranded DNA templates by purified VirD2 protein. Proc. Natl. Acad. Sci. USA (in press).
C. Koncz, N. Martini, L. Szabados, M. Hrouda, A. Bachmair and J. Schell (1993). Specialised vectors for gene tagging and expression studies. In: PLant Molecular Biology Manual, S. Gelvin and B. Schilperzoort (eds), Kluwer Academic, Dordrecht (in press).
J. Mathur, C. Koncz and L. Szabados (1994). A simple method for isolation, liquid culture, transformation and regeneration of *Arabidopsis thaliana* protoplasts. Plant Cell Report (submitted).
J.J.J. Ooms, S.M. de Bruijn and C.M. Karssen (1993). Acquisition of desiccation tolerance in seeds of *Arabidopsis thaliana*. In: Basic and Applied Aspects of Seed Biology, D. Come and F. Corbineau (eds), ASFIS Paris ISBN 2-9507351-1-8.
S.M. de Bruijn, J.J.J. Ooms, A.S. Basra, A.A.M. van Lammeren and D. Vreugdenhill (1993). Influence of abscisic acid on storage of lipids and carbohydrates in developing *Arabidopsis* seeds. In: Basic and Applied Aspects of Seed Biology, D. Come and F. Corbineau (eds), ASFIS Paris ISBN 2-9507351-1-8.
J.J.J. Ooms, J.M. Wilmer and C.M. Karssen (1994). Carbohydrates are not the sole factor determining desiccation tolerance in seeds of *Arabidopsis thaliana*. Physiol. Plant. (in press).
J.J.J. Ooms, R. van der Veen and C.M. Karssen (1994). Slow dehydration induces desiccation tolerance in seeds of desiccation intolerant ABA-deficient and ABA-insensitive mutants. (submitted).
K.M. Leon-Kloosterziel, C.J. Keijzer and M. Koornneef (1994). A seed shape mutant of *Arabidopsis*, which is integument development. Plant Cell. (in press).
K.J. Halliday, M. Koornneef and G.C. Whitelam (1994). Phytochrome B, and at least another phytochrome, mediate the accelerated flowering response of *Arabidopsis thaliana* L. to low red: far-red ratio. Plant Physiol. (in press).
M.C. Clarke, W. Wei and K. Lindsey (1992). High frequency transformation of *Arabidopsis thaliana* by *Agrobacterium tumefaciens*. Plant. Molec. Biol. Reporter **10**: 178-189.
K. Lindsey, W. Wei, M.C. Clarke, H.F. McArdle, L.M. Rooke and J.F. Topping (1993). Tagging genomic sequences that direct transgene expression by a activation of a promoter trap in plants. *Transgenic Research* **2**: 33-47.
K. Lindsey and J.F. Topping (1993). Embryogenesis: a question of pattern. *J. Exp. Bot* **44**: 359-374.
J.F. Topping, F. Agyeman, B. Henricot and K. Lindsey. Identification of molecular markers of embryogenesis in *Arabidopsis thaliana* by promoter trapping., (submitted)
C. Grevelding, V. Fantes, E. Kemper, J. Schell and R. Masterson (1993). Single-copy T-DNA insertions in *Arabidopsis* are the predominant form of integration in root derived transgenics, while multiple insertions are found in leaf-discs. Plant Molec. Biol. (in press).
M. Anderson and B. Mulligan (1993), Nottingham *Arabidopsis* Stock Centre Seed List. 120 pp.
M. Pages, M. Pla, J. Vilardell, A. Goday (1994). Water-stress responsive genes in maize. In Biotechnology for Aridland Plants. T.J. Mabry, H.T. Nguyen, R.A. Dixon, M.S. Bonness (eds) IC^2 The University of Texas at Austin, Texas, pp. 35-46.
M.M. Alba, C. Arenas, M. Pages. Subcellular localisation of the RNA-binding protein MA16 in maize and in transgenic *Arabidopsis* plants (in preparation).
M.G.M. Aarts, W. Dirkse, W.J. Stiekema and A. Pereira (1993). Transposon tagging of a male sterility gene in *Arabidopsis*, Nature, **363**: 715-717.
G. Cnops, B. den Boer, G. Engler, M. van Montagu and M. van Lijsebettens (1994). Towards the cloning of the TORNADO locus. Arch. Int. Physiol. Biochim. Biophys. (in press).
R. Schmidt, G. Cnops, I. Bancroft and C. Dean (1992). Construction of an overlapping YAC library of the *Arabidopsis thaliana* genome. Aust. J. Plant Physiol. **19**: 341-351.
R. Schmidt and C. Dean (1992). Physical mapping of the *Arabidopsis* genome. In Genome Analysis, vol. 4. K. Davies and S. Tilghman (eds). CSH publication 71-98.
I. Bancroft, L. Westphal, R. Schmidt and C. Dean (1992). PFGE-resolved RFLP analysis of the DNA of *Arabidopsis thaliana* using probes as large as whole YACs Nuc. Acid Res. **20**: 6201-6207.

R. Schmidt and C. Dean (1993). Towards construction of an overlapping YAC library of the *Arabidopsis thaliana* genome. Bio Essays **15**: 63-69.
M. Anderson and B. Mulligan (1992). Nottingham *Arabidopsis* Stock Centre Seed List.
M. Anderson and B. Mulligan (1992). *Arabidopsis* Mutant Collection. In: C.S. Koncz, M.-H. Chua and J. Schell (eds). Methods in *Arabidopsis* Research. World Scientific, Singapore, pp. 419-437.
D. Long and G. Coupland (1993). The maize transposable element system *Ac/Ds* as a mutagen in *Arabidopsis*: Identification of an *albino* mutation induced by *Ds* insertion (submitted).
D. Long and G. Coupland (1993). Analysis of the inheritance of transposed *Ds* elements in *Arabidopsis* after activation by a CaMV 35S promoter fusion to the *Ac* transposase gene (submitted).
C. Koncz, K. Németh, G.P. Rédei and J. Schell (1992). T-DNA insertional mutagenesis in *Arabidopsis*. Plant Mol. Biol. **20**: 963-976.
C. Koncz, N.-H. Chua and J. Schell (1992). Methods in *Arabidopsis* Research, World Scienfitic, Singapore.
C. Grevelding, D. Becker, R. Kunze, A. von Menges, V. Fantes, J. Schell and R. Masterson (1992). High rates of *Ac/Ds* germinal transposition in *Arabidopsis* suitable for gene isolation by insertional mutagenesis. Proc. Natl. Acad. Sci. USA, **89**: 6085-6089.
J. Putterill, F. Robson, K. Lee and G. Coupland (1993). Chromosome walking with YAC clones in *Arabidopsis*: isolation of 1700 kb of contiguous DNA on chromosome 5, including a 300 kb region containing the flowering-time gene *CO*. Mol. Gen. Genet. (in press).
C. Dean, I. Bancroft, A. Bhatt, J. Clarke, J. Chandler, G. Cnops, R. Ewing, E. Lawson, C. Lister, T. Page, R., Schmidt and L. Westphal (1992). Cloning genes involved in floral induction in *Arabidopsis thaliana*. The Flowering Newsletter **11**: 17-19.
I. Bancroft, L. Westphal, R. Schmidt and C. Dean (1992). PFGE-resolved RFLP analysis of the DNA of *Arabidopsis thaliana* using probes as large as whole YACs. Nuc. Acid Res. **20**: 6201-6207.
M. Raynal, L. Aspart, P. Gaubier, D. Depigny, F. Grellet and M. Delseny (1992). Gene expression during seed formation and maturation in Crucifereae. In: Reproductive Biology and Plant Breeding. Y. Dattee, C. Dumas and A. Gallais (eds), pp. 225-234.
J. Ooms, S. de Bruijn, C. Karssen (1992). Acquisition of desiccation tolerance in seeds of *Arabidopsis thaliana*. In: D. Come and F. Corbineau (eds), Basic and Applied Aspects of Seed Biology (in press).
S. de Bruijn, J. Ooms, A. Basra, A. van Lammeren, D. Vreugdenhil (1992). Influence of abscisic acid on storage of lipids and carbohydrates. In: D. Come and F. Corbineau (eds), Basic and Applied Aspects of Seed Biology (in press).
J. Ooms, J. Wilmer, C. Karssen. The role of carbohydrates in acquisition of desiccation tolerance in seeds of *Arabidopsis thaliana*, manuscript in preparation.
J. Giraudat, B.M. Hauge, C. Valon, J. Smalle, F. Parcy and H.M. Goodman (1992). Isolation of the *Arabidopsis* ABI3 gene by positional cloning. The Plant Cell **4**: 1251-1261.
M.C. Clarke, W. Wei and K. Lindsey (1992). High frequency transformation of *Arabidopsis thaliana* by *Agrobacterium tumefaciens*. Plant Molec. Biol. Reporter **10**: 178-179.
K. Lindsey, W. Wei, M.C. Clarke, H.F. McArdle, L.M. Rooke and J.F. Topping (1993). Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. Transgenic Research **2**: 33-47.
K. Lindsey and J.F. Topping (1993). Embryogenesis: a question of pattern. J. Exp. Bot. **44**: 359-374.

T-PROJECT

LIPASES

3-Dimensional structure and catalytic mechanism of 2-3 selected lipases of industrial relevance (BIOT CT-900181)

COORDINATOR:
L. THIM, Novo Nordisk A/S, Bagsværd, DK
PARTICIPANTS:
G.G. DODSON, University of York, York, GB
R. VERGER, C.N.R.S., Marseille, F

BACKGROUND INFORMATION

Partly as a result of the Biotechnology Action Programme (BAP), the first 3D-structure of a triglyceride lipase was published in Nature in 1990 (Brady, L. et al., A serine protease triad forms the catalytical centre of a triacylglycerol lipase, Nature **343** (1990) 767-770). This paper contained the background information for further studies of the two fungal lipases of the present project.

OBJECTIVES AND PRIMARY APPROACHES

The project was divided into two parts, one concerning the study of the two fungal lipases from *Rhizomucor Miehei* (RML) and *Humicola Lanuginosa* (HLL), respectively and one concerning the study of the mammalian pancreatic lipases from human (HPL), guinea pig (GPLRP2) and coypu (CoPLRP2). Recombinant lipase expression and site-directed mutagenesis have been performed at Novo Nordisk (Denmark), kinetic characterization of lipases and mutants at CNRS Marseilles (France) and X-ray crystallography at University of York (UK).

RESULTS AND DISCUSSION

A. Fungal lipases (RML and HLL)

By using the techniques of X-ray crystallography, we have determined the mechanism for 'interfacial activaton' in the lipases isolated from the filamentous fungi Humicola lanuginosa (HLL) and Rhizomucor miehei (RML). From previous structural studies, it was apparent that in both enzymes the active site contains a serine protease-like catalytic triad which is buried by a short amphipathic helix (referred to as the 'lid'). Through inhibitor and substrate analogue binding studies, we now have several X-ray structures of these lipases in interfacially activated states (i.e. the lid helix is displaced to one side, thereby exposing the active site). This conformational change generates a huge hydrophobic surface around the active centre, which is presumably stabilized by adsorption to a non-polar lipid interface. Moreover these structures demonstrate how the high energy tetrahedral transition state of the lipolytic mechanism is stabilized by an oxy-anion hole which is formed upon activation.
We have attempted to model triglyceride binding in HLL. Possible interactions between this lipase and its substrate have been analysed through molecular modelling and dynamics. These studies have enabled us to postulate a number of potential binding sites for fatty acid moieties of the substrate. This work has since been complemented by a parallel study in which HLL was crystallized in the presence of phospholipid/bile salt mixed micelles (NB. phospholipids are not hydrolysed by this enzyme). Close analysis of the active site region in this structure revealed that the lid was open and a phospholipid molecule could be modelled into electron density adjacent to the active serine (see figure 1).

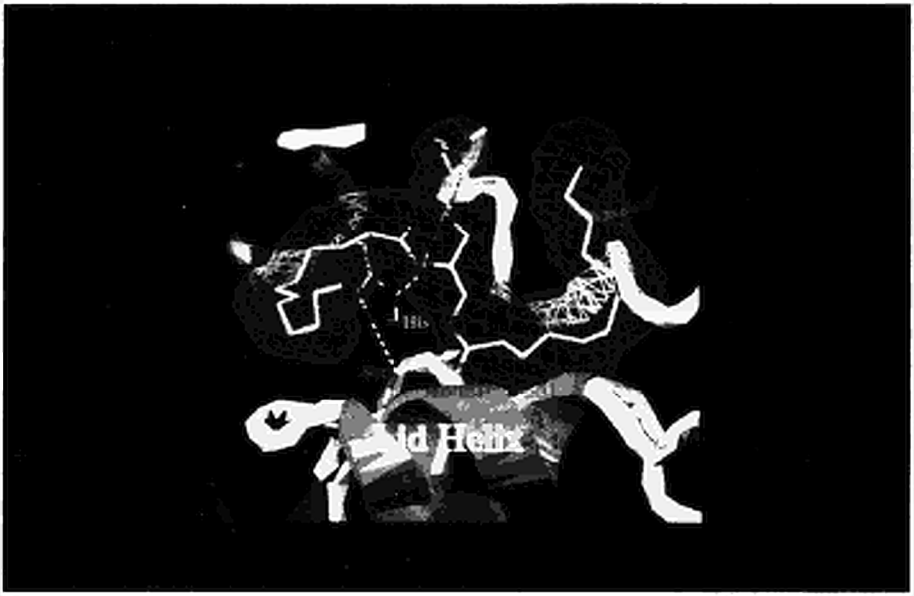


Fig. 1 shows a molecule of dilauryl phosphatidyl choline modelled into 3.2Å resolution positive difference electron density at the active site of Humicola lanuginosa lipase which is in an interfacially activated conformation. Although phospholipids are not substrates of this enzyme, the sn1 chain is bound in the hydrophobic groove which extends away from the active serine. The equivalent of the scissile ester bond of the true substrate is adjacent to the serine, and the direction of the nucleophilic attack is indicated by the arrow.

It is not obvious from the above model, why this substrate analogue is not hydrolysed, as it appears to be ideally placed over the active site with the equivalent of the scissile bond within range of the nucleophilic active serine. Work is in progress to clarify these observations by further X-ray analysis.

B. Mammalian pancreatic lipases (HPL, GPLRP2, CoPLRP2)

Guinea pig pancreatic lipase (GPLRP2) differs from other pancreatic lipases in that it is not interfacially activated, but displays normal Michaelis-Menten kinetics. Furthermore, its activity is unaffected by the presence of bile salts and/or colipase and it exhibits phospholipase A1 activity. The isolation, sequencing, cloning, expression and enzymatic characterization of the GPLRP2 has been carried out. The amino acid sequence of GPLRP2 is highly homologous to other known pancreatic lipases, with the exception that it carried a large deletion in the lid domain. In absence of useful X-ray data, GPLRP2 modelling was performed based on human pancreatic lipase (HPL) 3D structure. From this modelling, the environment of the catalytic centre appeared to be conserved overall, with the exception of the lid domain. We then proposed that the lid deletion in GPLRP2 was responsible for the anomalous behaviour of this enzyme, mainly the absence of interfacial activation and a high phospholipase A1 activity on all the phospholipids classes whereas classical pancreatic lipases such as HPL only hydrolyse negatively charged phospholipids with a very low turnover.

In search for similar enzymes in species closely related to guinea pig, we also studies the coypu pancreatic lipase related protein 2 (CoPLRP2). This pancreatic lipase possesses a full length lid domain but it shares the same kinetic properties as GPLRP2:

(1) absence of interfacial activation,

(2) inhibition by bile salts of the true lipase activity on insoluble triglycerides,

(3) no effect of colipase which normally counteracts bile salt inhibition, and

(4) high phospholipase activity.

Since both guinea pig and coypu pancreas produce a classical pancreatic lipase and no measurable phospholipase A_2 activity, we suggested that GPLRP2 and CoPLRP2 act as real phospholipases under physiological conditions, and belong to a pancreatic lipase subfamily (RP2) where all the enzymes share phospholipase activity and fulfil new biological functions when compared to the classical lipase. Since the size of the lid domain is different in GPLRP2 and CoPLRP2, we concluded that this domain was not essential for phospholipase activity. This could explain why this domain has been extensively mutated within the RP2 subfamily through evolution. The structural reason for phospholipase activity of RP2 lipase is, however, still unknown. The presence of a full-length lid domain in CoPLRP2 also led to the conclusion that the simple presence of a lid was not synonymous with interfacial activation, and the lid domain in CoPLRP2 is probably always found in an open conformation.

The structure-function relationships between the lid domain and the kinetic properties were investigated using site-directed mutagenesis and domain exchange between GPLRP2 and the classical human pancreatic lipase (HPL). The replacement of HPL lid by the mini-lid of GPLRP2 mainly resulted in impairing the specific activity of the chimera, but as suggested from GPLRP2 modelling, interfacial activation was lost probably because the active site became directly accessible. No phospholipase activity could be measured. On the other hand, the replacement of GPLRP2 mini-lid by HPL lid did not confer interfacial activation and did not suppress phospholipase activity. These results confirmed the independence of phospholipase activity with respect to the lid domain, and it also shows that the environment of the lid was important for interfacial activation.

Most of the recombinant pancreatic lipases and mutants characterized during this study were produced in insect cells using the Baculovirus expression system. This system was extremely efficient for production of small amounts of lipases for structure-function studies. In the case of HPL as a model enzyme, the expression level reached 40 mg of enzyme per liter of serum-free culture medium, and a single cationic exchange chromatography was sufficient to obtain a highly pure recombinant HPL with kinetic properties similar to those of the native enzyme purified from pancreatic juice.

PUBLICATIONS

Joint publications

Brzozowski, A.M., Derewenda, U., Derewenda, Z.S., Dodson, G.G., Lawson, D.M., Turkenburg, J.P., Bjorkling, F., Huge-Jensen, B., Patkar, S.A. & Thim, L. A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. Nature (1991) **351**, 491-494.

Hjorth, A., Carrière, F., Cudrey, C., Woldike, H., Boel, E., Lawson, D.M., Ferrato, F., Cambillau, C., Dodson, G.G., Thim, L. & Verger, R. A structural domain (the lid) found in pancreatic lipases is absent in the guinea pig (phospho)lipase. Biochemistry (1993), **32**, 4702-4707.
Thirstrup, K., Carrière, F., Hjorth, S., Rasmussen, P.B., Wöldike, H., Nielsen, P.F. and Thim, L. One-step purification and characterization of human pancreatic lipase expressed in insect cells. FEBS Lett. (1993) **327**, 79-84.
Carrière, F., Thirstrup, K., Hjorth, S. and Boel, E. Cloning of the classical guinea pig pancreatic lipase and comparison with the lipase related protein 2. FEBS Lett. (1994a) **338**, 63-68.
Thirstrup, K., Verger, R. and Carrière, F. Evidence for a pancreatic lipase subfamily with new kinetic properties. Biochemistry (1994) **33**, 2748-2756.
Carrière, F., Thirstrup, K., Boel, E., Verger, R. and Thim, L. Structure-function relationships in naturally occurring mutants of pancreatic lipase. Prot. Eng. (1994b) **7**, 563-569.
Thirstrup, K., Carrière, F., Hjorth, S., Rasmussen, P.B., Ladefoged, C., Thim, L. and Boel, E. Cloning and expression in insect cells of two pancreatic lipases and colipase from *Myocastor coypus*, (1995) (submitted).
Carrière, F., Thirstrup, K., Hjorth, S., Nielsen, P.F., Verger, R., Boel, E. and Thim, L. Investigation of pancreatic lipase structure-function relationships by domain exchange. (1995) (in preparation).
Verger, R., Ferrato, F., Carrière, F., Cudrey, C., Rugani, N., Gargouri, Y., Hjorth, A., Wöldike, H., Boel, E., Thim, L., Lawson, D.M., Dodson, G.G., van Tilbeurgh, H., Egloff, M.P. and Cambillau, C. Relationships between structures and kinetic properties of pancreatic lipases. Annals of the New York Academy of Sciences. Enzyme Engineering (Deauville, September 19-24, 1993).
Verger, R., Ferrato, F., Aoubala, M., de Caro, A., Ivanova, M., de la Fournière, L., Carrière, F., Cudrey, C., Rivière, C., Rugani, N., Gargouri, Y., Hjorth, A., Wöldike, H., Boel, E., Thim, L., Lawson, D.M., Dodson, G.G., van Tilbeurgh, H., Egloff, M.P. and Cambillau, C. Lipases and Esterases: reassessment of their classification in the light of their three dimensional structure. Proceedings of the NATO-ARW 'Esterases, Lipases, Phospholipases: From Structure to Clinical Significance' (University of Bordeaux II, September 22-24, 1993). Plenum Publishing Corporation. Editor: M.I. Mackness.
Verger, R., van Tilbeurgh, H., Ferrato, F., Rogalska, E., Aoubala, M., de Caro, A., Ivanova, M., de la Fournière, L., Rivière, C., Sarda, L. and Cambillau, C. Relationships between structures and kinetic properties of Lipases. Proceedings of the Sixth European Congress on Biotechnology (Florence — June 13-17, 1993). Elsevier. Eds: L. Alberghina, L. Frontali and P. Sensi (pp. 121-128).
Martinez, C., Nicolas, A., van Tilbeurgh, H., Egloff, M.P., Cudrey, C., Verger, R. and Cambillau, C. Cutinase, a lipolytic enzyme with a preformed oxyanion hole. Biochemistry (1994) **33**, 83-89.
Cudrey, C., van Tilbeurgh, H., Gargouri, Y. and Verger, R. Inactivation of pancreatic lipases by amphiphilic reagents (C12-TNB, THL). Dependence upon partitionning between micellar and oil phases. Biochemistry (1993) **32**, 13800-13808.
Ransac, S., Deveer, A.M.T.J., Rivière, C., Slotboom, A., Gancet, C., Verger, R. and de Haas, G.H. Competitive inhibition of lipolytic enzymes. V. A monolayer study using enantiomeric acylamino analogues of phospholipids as potent competitive inhibitors of porcine pancreatic phospholipase A_2. Biochim. Biophys. Acta. (1992) **1123**, 92-100.
Deveer, A.M.T.J., Franken, P.A., Dijkman, R., Meeldijk, J., Egmond, M.R., Verheij, H.M., Verger, R. and de Haas, G.H. Competitive inhibition of lipolytic enzymes. VII. The interaction of pancreatic phospholipase A_2 with micellar lipid-water interfaces of competitive inhibitors. Biochim. Biophys. Acta. (1992) **1125**, 73-81.
Ransac, S., Aarsman, A.J., van den Bosch, H., Gancet, C., de Haas, G.H. and Verger, R. Rat platelet phospholipase A_2. Kinetic characterization using the monomolecular film technique. Eur. J. Biochem. (1992) **204**, 793-797.
Deveer, A.M.T.J. den Ouden, A.T., Vincent, M., Gallay, J., Verger, R., Egmond, M.R., Verheij, H.M. and de Haas, G.H. Competitive inhibition of lipolytic enzymes. VIII: Inhibitor-induced aggregation of porcine pancreatic phospholipase A_2. Biochim. Biophys. Acta. (1992) **1126**, 95-104.
Van Tilbeurgh, H., Sarda, L., Verger, R. and Cambillau, C. Structure of the pancreatic lipase-procolipase complex. Nature (1992) **359**, 159-162.

Van Tilbeurgh, H. Gargouri, Y., Dezan, C., Egloff, M.P., Nésa, M.P., Rugani, N., Sarda, L., Verger, R. and Cambillau, C. Crystallization of pancreatic procolipase and of its complex with pancreatic lipase. J. Mol. Biol. (1993) **229**, 552-554.
Van Tilbeurgh, H., Egloff, M.P., Martinez, C., Rugani, N., Verger, R. and Cambillau, C. Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography. Nature (1993) **362**, 814-820.

Individual publications

Derewenda, U., Brzozowski, A.M., Lawson, D.M. and Derewenda, Z.S. Catalysis at the interface: the anatomy of a conformational change in triglyceride lipase. Biochemistry (1992) **31**, 1532-1541.
Lawson, D.M., Brzozowski, A.M. and Dodson, G.G. Lifting the lid off lipases. Current Biology (1992) **2**, 473-475.
Lawson, D.M., Brzozowski, A.M., Rety, S., Verma, C. and Dodson, G.G. Probing the nature of substrate binding in *Humicola lanuginosa* lipase through X-ray crystallography and intuitive modelling. Protein Engineering (1994) **7**, 543-550.
Marguet, F., Cudrey, C., Verger, R. and Buono, G. Digestive lipases: inactivation by phosphonates. Biochim. Biophys. Acta. (1993) **1210**, 157-166.
Laurent, S., Ivanova, M., Pioch, D., Graille, J. and Verger, R. Interactions between ß-cyclodextrin and insoluble glyceride monomolecular films at the argon/water interface. Application to lipase kinetics. Chem. Phys. Lipids. (1994) **70**, 35-42.
Bénicourt, C., Blanchard, C., Carrière, F., Verger, R. and Junien, J.L. Potential use of a recombinant dog gastric lipase as an enzymatic supplement to pancreatic extracts in cystic fibrosis. Clinical Ecology of Cystic Fibrosis (1993). Eds. H. Escobar, C.F. Baquero and L. Suarez.
Redgrave, T.G., Ivanova, M. and Verger, R. The condensing effects of egg lecithin and cholesterol on triolein monolayers are inhibited by substitution of one saturated acyl chain in the triglyceride. Biochim. Biophys. Acta. (1994) **1211**, 229-233.
Carrière, F., Gargouri, Y., Moreau, H., Ransac, S., Rogalska, E. and Verger, R. Gastric lipases: Cellular, Biochemical and Kinetic Aspects. In: Lipases: their structure, biochemistry and application. Wooley P. and Petersen S.B. eds. Cambridge University Press, Cambridge (1994) 181-205.
de la Fournière, L., Ivanova, M.G., Blond, J.P., Carrière, F. and Verger, R. Surface behaviour of human pancreatic and gastric lipases. Colloids and Surfaces B: Biointerfaces (1994) (in press).
Labourdenne, S., Gaudry-Rolland, N., Letellier, S., Lin, M., Cagna, A., Esposito, G., Verger, R. and Rivière, C. The oil-drop tensiometer: potential applications for studying the kinetics of (phospho)lipase action. Chem. Phys. Lipids. (1994) (in press).
Aoubala, M., Bonicel, J., Bénicourt, C., Verger, R. and de Caro, A. Tryptic cleavage of gastric lipases. Location of the single disulfide bridge. Biochim. Biophys. Acta. (1994) (in press).
Raneva, V., Ivanova, Tz., Verger, R. and Panaiotiv, I. Comparative kinetics of phospholipase A_2 action on liposomes and monolayers of phosphatidylcholine spread at the air-water interface. Colloids and Surfaces B: Biointerfaces (1994) (in press).
Gargouri, Y., Cudrey, C., Mejdoub, H. and Verger, R. Inactivation of human pancreatic lipase by 5-dodecyl dithio-2-nitrobenzoic acid. Eur. J. Biochem. (1992) **204**, 1063-1067.
Moreau, H., Abergel, C., Carrière, F., Ferrato, F., Fontecilla-Camps, J.C., Cambillau, C. and Verger, R. Isoform purification of gastric lipases: towards crystallization. J. Mol. Biol. (1992) **225**, 147-153.
Gargouri, Y., Ben Salah, A. and Verger, R. Importance de la lipase gastrique humaine dans la digestion des lipides alimentaires. Rev. Franç. Corps Gras. (1992) **7/8**, 207-213.
Rogalska, E., Ransac, S. and Verger, R. Controlling lipase stereoselectivity via surface pressure. J. Biol. Chem. (1993) **268**, 792-794.
Rogalska, E., Cudrey, C., Ferrato, F. and Verger, R. Stereoselective hydrolysis of triglycerides by animal and microbial lipases. Chirality (1993) **5**, 24-30.
Aoubala, M., Daniel, C., de Caro, A., Ivanova, M., Hirn, M., Sarda, L. and Verger, R. Epitope mapping and immunoactivation of human gastric lipase using five monoclonal antibodies. Eur. J. Biochem. (1993) **211**, 99-104.
Ivanova, M.G. Aoubala, M., de Caro, A., Daniel, C., Hirn, J. and Verger, R. A study on human gastric lipase and complexes with monoclonal antibodies using the monomolecular film technique. Colloids and Surfaces B: Biointerfaces. (1993) **1**, 17-22.

Carrière, F., Laugier, R. Barrowman, J.A., Douchet, I., Priymenko, N. and Verger, R. Gastric and pancreatic lipase levels during a test meal in dogs. Scand. J. Gastroenterol. (1993) **28**, 443-454.
Carrière, F., Barrowman, J.A., Verger, R. and Laugier, R. Secretion and contribution to lipolysis of digestive lipases during a test meal in humans: Gastric lipase remains active in the duodenum. Gastroenterology (1993) **105**, 876-888.
Ivanova, Tz., Raneva, V., Panaiotov, I. and Verger, R. Kinetics of spreading of DOPC liposomes at the air-water interface subjected to phospholipase A_2 hydrolysis. Colloïd Polym. (1993) **271**, 290-297.
Bourhim, N., Elkebbaj, M.S., Gargouri, Y., Giraud, P., Van Rietschoten, J., Oliver, C. and Verger, R. Penetration of phospholipid monolayers by opioids peptides and opiates drugs (agonists and antagonists). Colloïds and Surfaces B: Biointerfaces. (1993) **1**, 203-211.
Abousalham, A., Rivière, M., Teissere, M. and Verger, R. Improved purification and biochemical characterisation of phospholipase D from cabbage. Biochim. Biophys. Acta. (1993) **1158**, 1-7.
Aoubala, M., Douchet, I., Laugier, R., Hirn, M., Verger, R. and de Caro, A. Purification of human gastric lipase by immuno affinity and quantifiction of this enzyme in the duodenal contents using a new ELISA procedure. Biochim. Biophys. Acta. (1993) **1169**, 183-188.
Gargouri, Y., Bensalah, A., Douchet, I. and Verger, R. Kinetic behaviour of four pancreatic lipases using emulsions and monomolecular films of glycerides. Biochim. Biophys. Acta. (1993) (in press).

Patent applications

Patent No. WO 93/00426. Priority: 25.06.92. Inventors: A. Hjorth, H.F. Wöldike, L. Thim, E. Gormsen, R. Verger.
Verger, R., Lin, M., Piéroni, G., Nury, S. and Cagna, G. Procédé de dosage des lipases. Brevet J.N.P.I. N° FR 87 11098.
Moreau, H., Verger, R., Lecat, D. and Junien, J.L. Lipase and lipasic extracts, process of preparation, and therapeutic use. European Patent N° 261 016 (1988).
Richou, J., Grimaldi, M., Verger, R., Rivière, C., Bois, A.G. and Nury, S. Procédé d'analyse physico-chimique fondé sur le contrôle de tensions interfaciales et appareil correspondant. Brevet J.N.P.I. N° FR 91 03105 (1991).
Verger, R., Rivière, C., Esposito, G. and Cagna, A. Procédé de mesure de l'activité des phospholipases. Brevet I.N.P.I. N° 92 11474 (1992).

Characterization of lipases for industrial application (BIOT CT-900194)

COORDINATOR:
M.R. EGMOND, Unilever Res. Lab., Vlaardingen, NL
PARTICIPANTS:
A. CLEASBY, M.E.M. Noble, L. Johnson Univ. of Oxford, Oxford, GB
G.H. de HAAS, H.M. Verheij, Univ. of Utrecht, Utrecht, GB
R. VERGER, CBBM/CNRS, F

BACKGROUND INFORMATION

Lipases (triacylglycerol hydrolases, E.C. 3.1.1.3) are interesting for application in industrial products and processes due to the wide range of reactions catalysed by these enzymes. For example, conversions include modification of fats and oils giving rise to higher quality products having better chemical and physical properties. Another important application area is the Detergent Industry where lipases are applied as adjuvants in fatty stain removal from polluted surfaces such as e.g. textiles.

Although many lipases are relatively robust, successful application of these enzymes in industry requires low operation costs. As a consequence, lipases need to be produced at low cost and also need to be adapted to the application environment. This may be accomplished in a most straightforward way using protein engineering methods. However, this requires a thorough understanding of lipase structure and function. At the start of this project on lipases no structural information existed. Based upon amino acid sequence information very little homology was found for lipases from various sources. Is was considered of utmost importance to elucidate the structures of several lipases in order to build up knowledge concerning structure and function relationships by combination of structural and kinetic data obtained under strictly controlled conditions. Apart from these relations between kinetic properties and structural features, the structural data will also help us to solve problems encountered when lipases are applied in the presence of other enzymes, in particular proteolytic enzymes of the subtilisin family.

The lipase of choice belongs to the family of *Pseudomonas* lipases. In particular it was found that *Pseudomonas glumae* lipase, formerly called *Chromobacterium viscosum* lipase gives rise to excellent performance when applied in Detergent systems. Although this enzyme is quite stable in the presence of detergents, it was found that subtilisins rapidly inactivate the enzyme, particularly in the presence of anionic detergents like alkyl sulphates that are currently used in washing systems.

In a wider context it is important to understand what controls the stereo- or positional preference of lipases while acting on (e.g.) triacylglycerols. The lipolytic reactions are complicated by the fact that lipases recognize lipid-water interfaces, a property that discriminates lipases from the general class of esterases. This implies that stereoselectivity may take place either at the level of interaction with the lipid-water interface, or after binding to the interface while interacting with single substrate molecules, or during both processes.

OBJECTIVES AND PRIMARY APPROACHES

— Obtain structural information of the lipase from the species *Pseudomonas glumae* (PGL). The approach taken is to crystallize the purified enzyme allowing X-ray crystallographic analysis (carried out at the University of Oxford, GB).

— Gather information concerning mechanism of action, substrate specificity and inhibitory compounds. Special triglyceride substrates and substrate analogues are made to elucidate stereo- and positional specificities of PGL (Utrecht, NL and Marseille, F).

— Generate tailor-made lipases for application in Detergent Industry also using understanding of structure-function relationships.

RESULTS AND DISCUSSION

Pseudomonas glumae *structural properties*

Pseudomonas glumae lipase contains 319 amino acids giving a calculated molecular mass of 33.1 kDa, which is close to the molecular mass of 33 kDa observed on SDS gels. There are 22 acidic (17 asp, 5 glu) and 17 basic amino acid residues (6 lys, 11 arg) yielding a calculated iso-electric point around 6.0, which differs from the observed value around 7.4. This may either indicate that about 4 or 5 acidic groups are shielded from solvent water and/or do not titrate. An alternative or additional explanation may be that the enzyme strongly binds one or more calcium ions for which acidic side chain groups provide liganding carboxylate residues.

The titration curve calculated from the primary structure is shown above in Fig. 1. This figure shows that at pH 7.4 the overall charge of the enzyme is about −5 when all acidic groups are titrating normally.

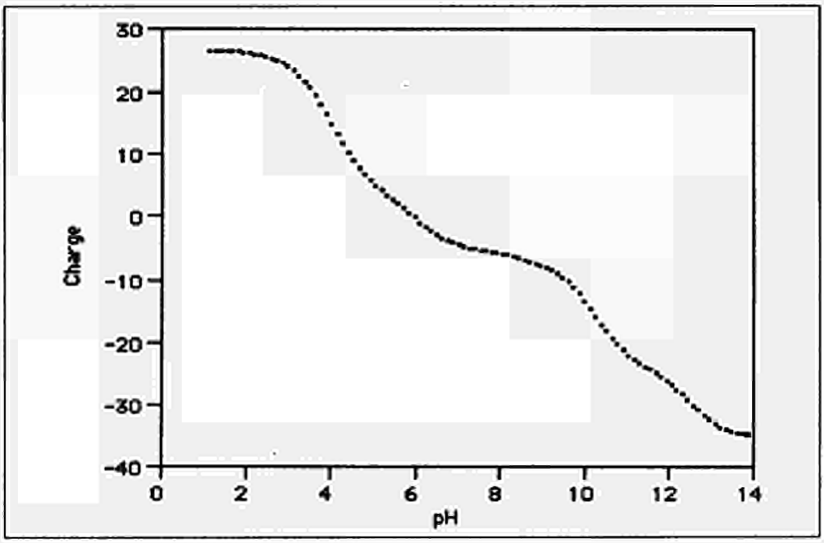


Fig. 1. Calculated curve derived from the amino acid composition of Pseudomonas glumae lipase.

The discrepancy between calculated and observed charge was solved to a large extent when the 3 dimensional structure was obtained using Xray crystallography. A typical view of the backbone structure of *Pseudomonas glumae lipase is shown in Figure 2.*

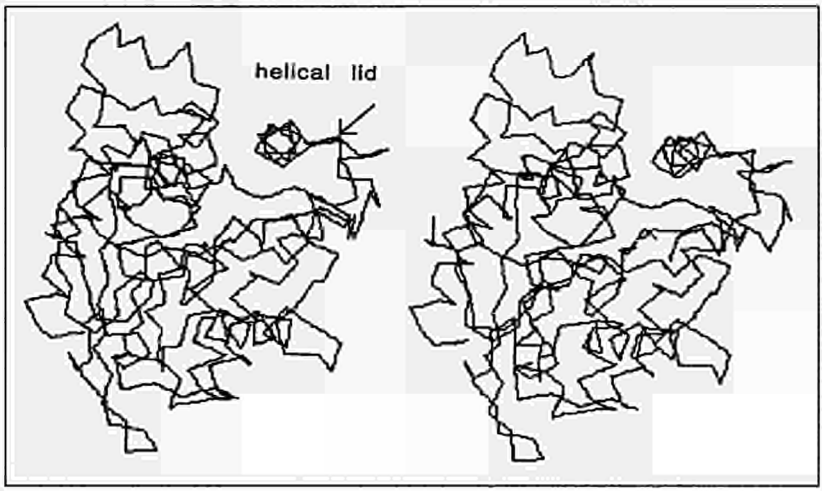


Figure 2. Stereo picture of Pseudomonas glumae lipase (PGL) backbone structure obtained from X-ray crystallographic analysis. The arrow indicates the proteolytic splitting site found.

As was found previously for several other lipase structures the active site residues of the crystallized protein are hidden from solvent water. For PGL the active site triad was found to consist of Ser87, His285 and Asp263. The first two of these residues had previously been assigned by protein engineering studies. Replacement of Ser87 and His 285 by Ala residues, respectively, yielded inactive enzyme. Protein engineering studies aiming at identification of the acidic residue in the active side triad pointed to Asp241 rather than Asp263 because lipolytic activity was lost in Asp241Ala, while retained to some extent (10% w.r.t. wild-type) in the Asp263Ala variant. This result demonstrates that conclusions from such replacement studies have to be taken with care. However, it also shows the importance of Asp241. The structural information provides a good explanation for this and may also give an explanation for residual activity of the Asp263Ala variant (see below).

In all *Pseudomonas* lipases and perhaps also in e.g. *Bacillus* and *Staphylococcal* lipases the recently identified calcium binding site seems to be present that will give rise to strengthening of the structure close to the active site. In PGL the calcium ion is liganded to the fully conserved side chain carboxylates of residues Asp241 anbd Asp287 and backbone oxygens from residues Gln291 and Val295. Already the slightest change in this area will give rise to conformational changes and loss in activity and/or protein stability.

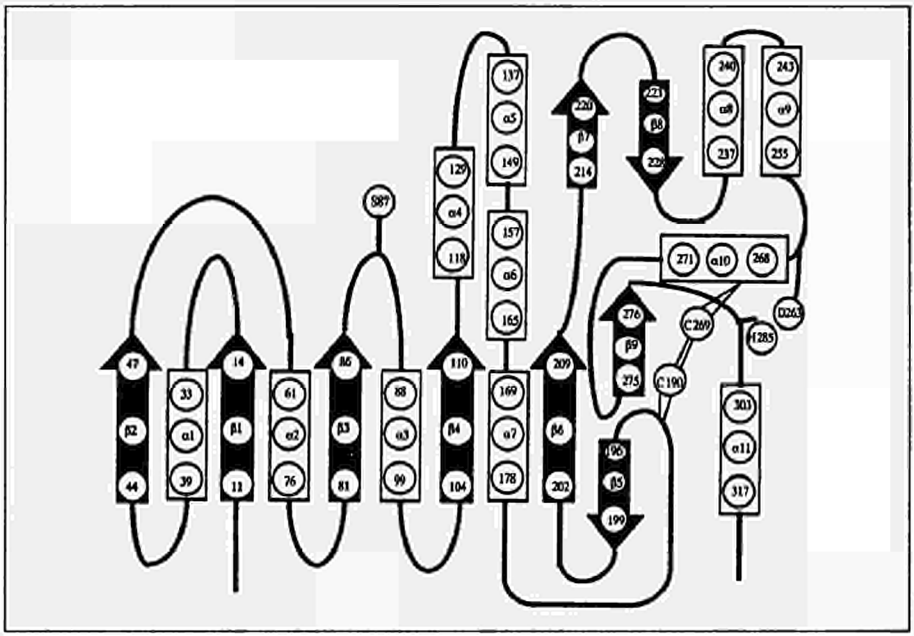


Figure 3. Secondary structure of Pseudomonas glumae as derived from the X-ray structure.

The strategic location of this calcium site may further be derived from its proximity to the active site His285 residue and another buried acidic residue Glu288. This latter residue is unique, as it is located within hydrogen bonding distance from the active site Asp263 residue. Possibly an important part is played by residue Glu288 as it provides the enzymes with an alternative acidic residue in the catalytic triad thus allowing the Aps263Ala variant to retain some residual activity. Additional mutation of Glu288 most likely will provide further proof for this hypothesis.

Figure 3 shows an overview of the secondary structure elements as derived from the tertiary structure of PGL.

The so-called hydrolase fold in this protein is built up from strands 1 through 4, strand 6 and the two small strands 5 and 9, sandwiched by helices 1 through 3, and also helices 7 and 11.

The part of the protein most likely involved in interaction with the substrate-water interface contains helices 4 through 6 and 8 through 10 as well as beta strands 7 and 8. Comparison of PGL with other *Pseudomonas* lipases tells us that beta strands 7 and 8 are unique for PGL and e.g. for *Pseudomonas cepacia* lipase, but lack in e.g. *Pseudomonas aeruginosa* lipase. The single disulphide bond connects cystein residues 190 and 269 respectively.

Proteolytic sensitive site

Limited proteolytic breakdown of wild-type *Pseudomonas glumae* lipase (PGL) by a subtilisin currently used in detergent systems (Savinase™ obtained from Novo Nordisk, DK) exclusively gives rise to only two fragments as judged from SDS-PAGE electrophoresis. Analysis of the smaller fragment (approx. 15 kDa) yielded

the N-terminal sequence of the wild type enzyme, whereas the larger fragment (approx. 18 kDa) starts with the sequence His-Asn-Thr-Asp-Gln-Asp-Ala. No other sequences were identified in this fragment. From this result it was concluded that PGL is split exclusively between residues Ser153 and His154. As this work was carried out before tertiary structure information was available, only limited conclusions could be drawn at that time concerning the actual splitting site. Secondary structure prediction methods based on the amino acid sequence indicated that the cleavage site is located in a non-structured flexible region of the protein (residues 151-157) flanked by a beta strand and helix at the N- and C-terminal sides, respectively.

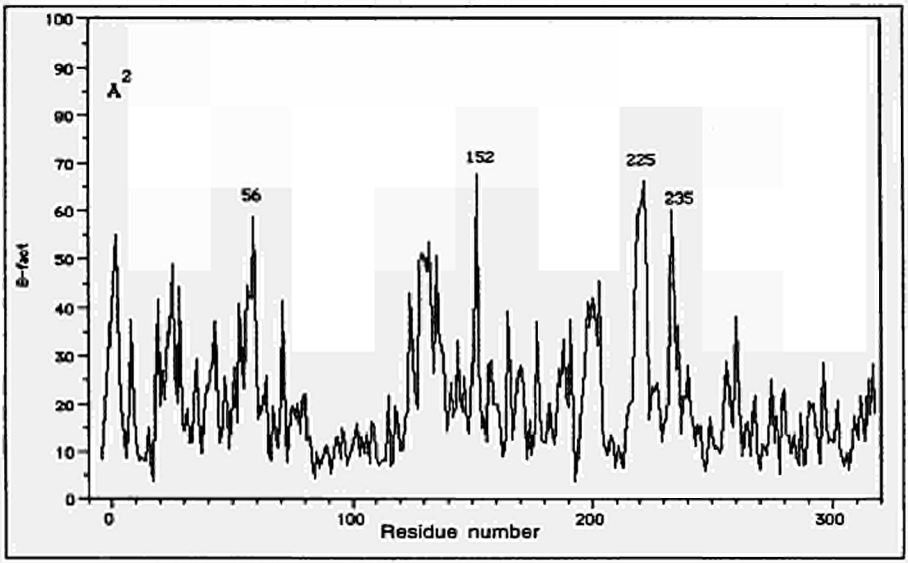


Figure 4. B factors as derived from the X-ray structure of PGL for backbone atoms only.

As is shown in Figure 2 the identification of the flexible peptide region (indicated by the arrow) was confirmed by X-ray analysis of PGL. However, this region is actually flanked by two helical structures running from residues 137-149 and 157-165, respectively. The sequence of the flexible peptide region around the splitting site reads: Val150-Ser151-Ser152-Ser153-His154-Asn155-Thr156. A hydrophobic residue at the P4-site (according to Schechter and Berger notation, here a Val residue at 150 is known to be preferred by the subtilisin protease Savinase.

Backbone flexibility as derived from crystallographic B-factors is indicated in Figure 4.

This Figure shows those areas of the protein backbone having highest structural uncertainty due to structural flexibility. The observed proteolytic splitting site is close to the site (residue 152) showing a very high B-factor indeed. Furthermore, structural variability has been observed for the four PGL molecules in the crystallographic asymmetric unit, particularly in one of the arms of helix 5. It may be observed from Figure 4 that also other regions in the protein show high B-factors, while no additional subtilisin-type splitting sites have been found. One such area is located near the two solvent-exposed antiparallel beta strands 7 and 8 (see Fig. 3

for overview of secondary structure layout). An intrinsic protease activity in *Pseudomonas* has been identified splitting the PGL sequence also between residues 219-220, and 220-221, respectively (data not shown).

It should be mentioned that the crystal structure of PGL has been obtained in the absence of an oil-water interface. While the X-ray data show that the active site region is deeply buried in the interior of PGL under these conditions, it is likely that the active site opens up in the presence of an oil water interface. This conformational change most likely involves helix 5 and its flanking regions. It might therefore be interesting to investigate whether the presence of an oil-water interface will affect the proteolytic susceptibility.

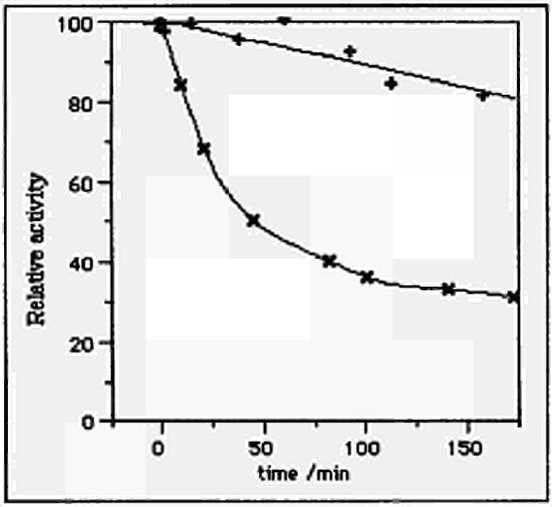


Figure 5. Relative lipolytic actiby in the presence of Savinase (pH 9) with (×) and without (+) micellar detergent.

Figure 5 shows that PGL is quickly degraded by Savinase in the presence of a detergent interface, while the enzyme is more stable without micellar detergent. This experiment suggests that the detergent interface triggers a conformational change leading to increased susceptibility to proteolytic attack. Apparently, this micellar interface is unable to block Savinase action indicating that while PGL is bound to micelles the arm of the helical flap is still exposed to solvent.

In summary, the present studies on PGL as a model of *Pseudomonas*-type lipases have given us some insight concerning proteolytic sensitivity being linked to essential flexibility in the lipase needed to allow access of substrate molecules to the active site triad. Although the observed calcium ion in PGL may not be catalytically important, its proximity to the active site at least points to involvement in retaining structural integrity close to the active site.

Substrate specificity

Kinetic analysis of lipolytic action is complicated due to the fact that the triacylglycerol substrates are poorly soluble or insoluble in water. In order to dissolve the lipid substrates this requires in most cases that lipases are investigated using 'inert' lipid-water interfaces. An additional problem then arises to discriminate enzyme binding to the interface from binding of single substrate molecules to the active site of the enzyme. In both of these processes substrate preference or specificity may play a role. Furthermore, for natural triacylglycerol substrates more than one fatty acid moiety is released, while acyl migration in diacyl- and monoacyl glycerol substrates may further complicate kinetic analyses.

For this reason optically pure pseudoglycerides were prepared in this project and investigated kinetically using well-defined detergent interfaces. In particular, the synthesis and purification has been completed for all six enantiometric triacyl- or

diacylglycerol analogues containing an O-methyl ether-, a decanoyl ester-, and a decanoyl amide function, respectively, at either position of 1-aminoglycerol and 2-aminoglycerol.

Enantiomeric preference is derived from the kinetic behaviour of PGL on these substrates dissolved in a micellar detergent interface at a low molar ratio relative to detergent. This interface is present at saturating amounts relative to the enzyme, thereby excluding effects on different interfacial properties of the substrates investigated.

Although these synthetic compounds differ from the natural substrates while being close in chemical behaviour, it turned out that the same stereopreference was obtained, i.e. the acyl chain at the sn-3 position was preferentially hydrolysed as compared to the sn-1 position. Previous work on natural substrates has indicated that this stereopreference is retained in diacyl-glycerols as well.

Interestingly, the pseudoglyceride substrates with an hydrolysable acyl chain at the 2 position turned out to be very good substrates too. Particularly the R-isomer is rapidly converted, whereas the S-isomer is an extremely poor substrate.

PUBLICATIONS

Joint publications

M.E.M. Noble, A. Cleasby, L.N. Johnson, M.R. Egmond, L.G.J. Frenken (1993). FEBS Lett. **331**: 123-128.

M.E.M. Noble, A. Cleasby, L.N. Johnson, M.R. Egmond, L.G.J. Frenken (1994). Protein Engineering **7**: 559-562.

A.M.T.J. Deveer, P.A. Franken, R. Dijkman, J. Meeldijk, M.R. Egmond, H.M. Verheij, R. Verger and G.H. de Haas (1992). Biochim. Biophys. Acta **1125**, 73-81.

A.M.T.J. Deveer, A.T. den Ouden, M. Vincent, J. Gallay, R. Verger, M.R. Egmond, H.M. Verheij and G.H. de Haas (1992). Biochi. Biophys. Acta **1126**, 95-104.

A.M.T.J. Deveer (1992). Thesis University of Utrecht, 'Mechanism of activation of lipolytic enzymes'.

Individual publications

L.G.J. Frenken (1993). Thesis University of Utrecht, *'Pseudomonas glumae* lipase: Characterization, biogenesis and protein engineering'.

L.G.J. Frenken, J.W. Bos, C. Visser, W. Muller, J. Tommassen, C.T. Verrips (1993). Molecular Microbiology **9**: 579-589.

L.G.J. Frenken, M.R. Egmond, A.M. Batenburg, C.T. Verrips (1993). Protein Engineering **6**: 637-642.

Exploring the structure-function relationship of *Pseudomonas* and *Bacillus* lipase (BIOT CT-910272)

COORDINATOR:
(1) O. MISSET, Gist-Brocades, Delft, NL
PARTICIPANTS:
(2) U. WINKLER, Ruhr Universität Bochum, Bochum, D
(3) C. COLSON, Univ. Cathol. Louvain-la-Neuve, Louvain-la-Neuve, B
(4) B.W. DIJKSTRA, Rijksuniversiteit Groningen, Groningen, NL

BACKGROUND INFORMATION

Lipases are unique in the sense that they are active on water/oil interfaces. Knowledge of their three-dimensional structures is essential for understanding the mechanism of action. The present project focused on the determination of the structures of the lipases from *P. aeroginosa* and *B. subtilis*.

OBJECTIVES AND PRIMARY APPROACHES

Our objective was the determination of at least one 3D-structure by X-ray crystallography. Until the succesful completion of this, reliable models were to be obtained computer modelling.
The construction of overproducing strains and the development of lipase purification protocols should result in large amounts of pure lipase. Subsequent crystallization experiments exploring a large variety of conditions should give crystals necessary for X-ray analysis. The role of specific amino acids would be examined by protein engineering.

RESULTS AND DISCUSSION

Fermentation and Purification

Since the wild type *B. subtilis* strain produced only very limited amounts of extracellular lipase ($\approx$ 1 mg/l culture medium), we had to construct an overproducing strain. This strain allowed for the purification of more than 100 mg of pure lipase. Table I-A gives a summary of the fermentation and the purification protocol.
For *P. aeruginosa* we had to adapt the fermentation procedure in order to optimize the lipase yield. Growing *P. aeruginosa* in Erlenmeyer flasks did not give high yields of lipase. Only fed-batch fermentation in 10 liter fermentors proved to result in reasonable expression levels. However, it was observed that the largest part of the lipase remained bound to the bacterial cells. This part could be solubilized by treatment of the cells with the non-ionic detergent Triton X-100. The purification procedure, involving preparative isoelectric focusing as a major step, resulted in lipase which was pure on the basis of protein content, but still contained lipopolysaccharides (LPS) which are present in the culture medium forming lipase-LPS micelles.

Recently, we managed to overexpress the lipase gene *lipA* in both *E. coli* and *P. aeruginosa* with the latter strain producing in Erlenmeyer flasks about 1000-fold the amount of enzymatically active lipase as compared to the wild type strain. At present we are in the process of testing the growth conditions for this overexpressing strain in 10 liter fed-batch fermentors.

Table I

Parameter	*P. aeruginosa*	*B. subtilis*
I-A: Fermentation and purification		
Fermentation		
strain	PAC1R, wild type	overproducing
device	10 L-fermentors	Erlenmeyer flask
localization of lipase	>90% cell bound	supernatant
expression level	20 — 50 mg/L	≈ 50mg/L
Purification		
step 1	Triton solubilization	$(NH_4)_2SO_4$-precipitation
step 2	Gel filtration	Phenyl-Sepharose
step 3	Isoelectric focusing	Hydroxyapatite
yield	10-15%	≈13%*
total mgs	≈ 80 mg	≈130 mg
purity	high, residual LPS	high
I-B: Biochemical properties		
# subunits	1	1
# amino acids	285	181
molecular weight	30,149	19,348
isoelectric point	5.9	9.9
# cysteines	2	0
disulfide bridge	yes	Not applicable
I-C: Kinetic properties		
Specific activity (units/mg)		
para-nitrophenyl-palmitate	45,600	160
triolein	22,800	Not Determined
Interfacial activation	no	no
Positional specificity	unspecific	1,3
pH-optimum	8	10

* *The yield has recently been improved towards ≈50%*

Biochemical properties

In Table I-B, the biochemical properties for both lipases have been summarized. Both enzymes are monomeric proteins with molecular weights of 30,149 (285 amino acids) and 19,348 (181 amino acids) Dalton for the *P. aeruginosa* and *B. subtilis* lipase, respectively, as was deduced from the gene-derived amino acid sequence. The lipase from *B. subtilis* is the smallest lipase known sofar. Whereas the *B. subtilis* lipase does not contain cysteines, the two cysteines in the *P. aeruginosa* lipase form a disulfide bridge.

Kinetic properties

In Table I-C, the major kinetic properties of both lipases are summarized. The specific activity of the *P. aeruginosa* lipase (measured at 37°C) is comparable to other lipases. On the contrary, the *Bacillus* lipase is much less active on *para*-nitrophenylpalmitate (PNPP). Surprisingly, although both enzymes are active on emulsified and non-emulsified substrates, they do not show interfacial activation with the short chain substrates triacetin and tripropionin.

Crystallization

The purified *P. aeroginosa* lipase was subjected to a large variety of crystallization conditions. Sofar, no crystals of sufficient size and quality have been obtained. An explanation for this unfortunate lack of success may the very hydrophobic nature of the lipase on the one hand and/or the presence of LPS on the other hand.

Table II
Crystallization and X-ray data for the *Bacillus subtilis* lipase

Crystallization conditions	
Protein concentration	6 mg/ml PEG-4000, 38-45% (w/v)
Precipitant	Na_2SO_4, 10-25 mM
Buffer	100 mM Ethanolamine, pH 9-10
Additive	0.7% n-octyl-β-D-glucoside
X-ray data	
Space group	Monoclinic C2
Cell dimensions	a = 121.5 Å; b = 93.2 Å; c = 81.0 Å α = γ = 90°; β = 110.67°
# molecules/AU*	4
Solvent content	57%
V_M	2.86 Å3/Da
Maximum resolution	≈ 2.5 Å
Native data set (from 2 crystals)	
Crystal dimensions	0.30 × 0.04 × 0.04 mm^3 (rods)
# observed reflections	134,197
# unique reflections	29,045
R_{sym}	9.8% (from 39.7 to 2.5 Å)
Completeness	99.7% (at 2.5 Å resolution)

* *AU = Asymmetric Unit*

B. subtilis lipase gave good quality crystals (see Table II. A complete native data set has been obtained from 2 crystals at the EMBL outstation at the DESY synchrotron in Hamburg. The structure elucidation was attempted by molecular replacement using either the X-ray structure of cutinase or the X-ray structure of the *P. glumae* lipase as a starting model. Although the rotation functions gave interpretable peaks, the translation function could not confirm the solutions obtained from the rotation function. Therefore, much effort was put into the search for heavy atom derivatives. Using 35 compounds, 47 different conditions were tested but, unfortunately, no good derivatives have been found until now.

Protein engineering of *B. subtilis* lipase

Table III summarizes the results of the various mutations as well as the reason why these have been performed.

Table III

Mutation	*Purpose*	*Result*
A75G	Restoration of the 'concensus sequence' Gly-Xxx-Ser-Xxx-Gly	64% of wild type activity less stable than wild type
S77C	Active site Serine → Cystein Serving as a binding site for heavy metal ions during crystallization	Inactive mutant Cys sensitive to oxidation 10 mg purified
H152N	Putative active site histidine	Same activity as wild type: no member of catalytic triad
H156N	Putative active site histidine	No activity: member of catalytic triad?
D118A	Putative active site aspartic acid	Same activity as wild type: no member of catalytic triad
D133N	Putative active site aspartic acid	No activity: member of catalytic triad?

Mutant A75G

Substitution of Ala75 by glycine did not result in improved lipase activity. The purified mutant revealed 64% of the wild type activity and appeared less stable as deduced from thermostability and activity tests.

Mutant S77C

This mutant was overproduced at the same level as the wild type enzyme, however, no activity could be detected in the fermentation broth. From this it can be concluded that Ser77 is the active site serine. A total of 10 mgs were purified via a slightly modified purification procedure, but the major part appeared oxidized and was therefore not suitable for the crystallization studies.

Mutants H152N and H156N

Both putative catalytic triad histidines were replaced by asparagine residues and both were overproduced at the same level as the wild type. While mutant H152N remained fully active, no lipase activity was detected in the growth medium of the H156N mutant which suggests that His156 belongs to the catalytic triad.

Mutants D118A and D133N

Both putative catalytic triad aspartic acid residues were replaced: the first one by alanine and the second by the isosteric asparagine. Both were expressed at wild type level and only D118A was fully active while D133N was inactive. This indicates that the latter belongs to the catalytic triad.

Model building of *P. aeroginosa* lipase

We built a model for the 3D-structure of the lipase from *P. aeroginosa* based on an alignment of 50 lipase/hydrolase amino acid sequences as well as several experimentally determined X-ray structures. Comparison of this model with the recently determined X-ray structure of the closely related *P. glumea* lipase (43% amino acid homology) revealed that the amino acids belonging to the catalytic triad (Ser^{82}, Asp^{229}, His^{251}) were correctly predicted as well as the majority of the secondary structural elements of the *Pseudomonas* enzyme. This work demonstrated that it is possible to correctly build the structure of a protein core even in the absence of obvious sequence homology with a functionally similar protein of known 3D-structure. However, several homologous protein sequences should be known to determine where the conserved parts in the sequences occur. These conserved parts can then be assigned to secondary structural elements of the known 3-D structure. Such models can than advantegeously be used for a preliminary search for mutagenesis sites.
Later the model was refined using the X-ray coordinates of the *P. glumae* lipase.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

(1) Overproducing strains of *B. subtilis* and *P. aeroginosa*.

(2) A refined and reliable 3D-model of the *P. aeroginosa* lipase.

(3) *B. subtilis* lipase crystals of sufficient quality for X-ray determination.

(4) Identification of the catalytic triad in the *Bacillus* lipase by site-directed mutagenesis.

MAJOR COOPERATIVE LINKS

The genetics and biochemistry of both lipases were carried out in the laboratories of '2' (*Bacillus* lipase) and '3' (*Pseudomonas*) who supplied '4' with pure enzyme for the crystallization and X-ray diffraction studies. '1' carried out the fermentations and supplied cell free supernatants to '2' and '3'. Nine meetings took place in Delft, Louvain-la-Neuve, Bochum and Groningen.
Additional links existed with various members of the other 4 project teams of the T-project.

PUBLICATIONS

Joint publications
Jaeger, K.-E., Ransac. S., Koch, H.B., Ferrato, F., and Dijkstra, B.W. (1993), FEBS Letters, **332**, 143-149
Ransac, S., Blaauw, M., Lesuisse, E., Schanck, K., Colson, C. and Dijkstra, B.W. (1994), J. Mol. Biol. **238**, 857-859
Misset, O., Gerritse, G., Jaeger, K.-E., Winkler, U., Colson, C., Schanck, K., Lesuisse, E., Dartois, V., Blaauw, M., Ransac, S. and Dijkstra, B.W (1994), Protein Engineering, 7, 523-529
Jaeger, K.-E., Ransac, S., Dijkstra, B.W., Colson, C., Heuvel, M. van, Misset, O. (1994), FEMS Microbiol. Rev. (in press)

Individual publications
Dartois, V., Baulard, A., Schanck, K. and Colson, C. (1992) Biochem. Biophys. Acta, **11311**, 253-260
Lesuisse, E., Schanck, K. and Colson, C. (1993), Eur. J. Biochem., **216**, 155-160
Dartois, V., Coppee, J.-Y, Colson, C. and Baulard, A. (1994) Appl. Environ. Microbiol (in press)

Triacylglycerol lipases and colipases of the digestive tract and exocellular fungal lipases: structure, interfacial binding and catalysis (BIOT CT-910274)

COORDINATOR:
R. VERGER, CNRS, Marseille, F
SUBCOORDINATORS:
A: L. THIM, Novo Nordisk, Bagsværd, DK
B: M.R. EGMOND, Unilever Research Laboratorium, Vlaardingen, NL
C: O. MISSET, Royal Gist Brocades NV, Delft, NL
D: R. VERGER, CNRS, Marseille, F
E: R.D. SCHMIDT, GBF, Braunschweig, D

BACKGROUND INFORMATION

The present report makes a fruitful use of two of the successful achievements of the BRIDGE-project: the 3-dimensional structures of cutinase as well as the Pancreatic Lipase/Colipase complex in its open and closed forms, that have been solved for the first time, with the presence of a phospholipid molecule clearly identifiable in the active site cleft of the Pancreatic Lipase.

OBJECTIVES AND PRIMARY APPROACHES

For many years, lipases have been attracting the attention of scientists in academic and industrial circles. One of the most intriguing and unique features of lipases is the way in which they are 'activated' by interfaces. The aim of this proposal is to elucidate of

(i) the structure;
(ii) the interfacial binding;
(iii) the catalytic mechanism of triacylglycerol lipases and colipases of the digestive tract and exocellular fungal lipase (cutinase).

Cutinase was choosen because of potential industrial applications. There is indirect evidence that colipase activates the pancreatic lipase-catalyzed interfacial hydrolysis by anchoring lipase to its substrate in the presence of bile salts. The colipase/lipase/interface is a challenging system of lipolytic regulation via protein/lipid and protein/protein interactions.

RESULTS AND DISCUSSION

1. We established the 2.4 Å resolution structure of the pancreatic lipase colipase complex covalently inhibited at the active site serine. Evidence has been showed for a pancreatic lipase subfamily with new kinetic properties. The structure-function relationships in naturally occuring mutants of pancreatic lipase and the kinetic behaviour of five pancreatic lipases using emulsions and monomolecular films of synthetic glycerides have been studied. We have also studied the inactivation components (C12-TNB, THL) and also dependence upon partitioning. We achieved the production of monoclonal antibodies to human pancreatic procolipase and the characterization of antibodies by competitive binding studies. We were able to identify a neutralizing antibody which recognizes a conformational epitope related to colipase activity.

2. The structure in solution of porcine pancreatic procolipase was determined form ^{1}H homonuclear 2D- and 3D-NMR experiments at 500- and 600-MHz. A protein fold element comprising a triple-strand beta-sheet and two disulphide bonds was recognized in procolipase and other unrelated proteins. Lipid binding and activating properties of porcine pancreatic colipase were showed to depend upon the proteolytic cleavage at the Ile_{79}-Thr_{80} bond.

3. A molecular model of lipoprotein lipase has been built based on the pancreatic lipase X-ray structure and the consequences for heparin binding and catalysis were proposed.

4. Human gastric lipase was purified by immunoaffinity and quantited in the duodenal contents using a new ELISA procedure. Gastric lipases were submitted to tryptic cleavage and a single disulfide bridge was positionned.

5. The interactions between β-cyclodextrin and insoluble glyceride monomolecular films at the argon/water interface have been investigated and applied to lipase kinetics studies. We studied the surface behaviour of human pancreatic and gastric lipases and we investigated the potential applications of the oil-drop tensiometer for studying the kinetics of (phospho)lipase action.

6. A comparative study on the competitive inhibition of pancreatic phospholipases A_2 from different sources by (R)-2-acylamino phospholipid analogues led us to a further delineation of the active site of pancreatic phospholipases A_2 from pig, ox and horse by comparing the inhibitory power of a number of (R)-2-acylamino phospholipid analogues.

7. The inactivation of *Staphilococcus hyicus* lipase by hexadecylsulfonyl fluoride has evidenced an active site serine for this enzyme. The inactivation of digestive lipases by phosphonates has been also investigated.

8. Cutinase, a lipolytic enzyme with a preformed oxyanion hole. The cutinase partitioning in PEG/potassium phosphate aqueous two-phase systems had been defined and we observed the ester synthesis by az recombinant cutinase in reversed micelles of a nautral phospholipid. Glycerine synthesis could be catalysed by cutinase using the monomolecular film technique.

9. Purification of Lipases from *Chromobacterium viscosum* raised the question of the existence of a high molecular weight lipase. Studies on the extraction and back-extraction of a recombinant cutinase in a reversed micellar system have been also performed.

10. An ultrafiltration membrane bioreactor for the lipolysis of olive oil in a reversed micellar media has been deviced. The triglycerides hydrolysis and the stability of recombinant cutinase from *Fusarium sonai* inAOT-isooctane reversed micelles have been studied. Polyglycerol-fatty acid esters have been enzymatically synthezied in a solvent-free system. Lipolysis in reversed micellar system was modeled both in conventional bath reactors and membrane reactors. The gas/solid reaction system, with a controlled water activity, was developed in a continuous process.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

1. The solution structure of procolipase has been determined by ^{1}H nuclear magnetic resonance (NMR).
2. The structures of the human pancreatic lipase/porcine procolipase complex, covalently inhibited at the active site serine by both enantiomers of C11 phosphonate compounds, have been established at 2.4 Å resolution.
3. Structure-function relationships in naturally occurring mutants of pancreatic lipase have been established.
4. Thirty five point mutants and complexes with inhibitors of cutinase have been studied by X-ray diffraction.
5. Spectroscopic characterization based on fluorescence of a recombinant cutinase encapsulated in reverse micelles.
6. Glyceride synthesis catalyzed by cutinase using the monomolecular film technique.
7. The solid-gas bioreactor as well as NMR spectroscopy are well suited techniques to study the role of water in lipase catalyzed reactions. The optimal activity and stability of cutinase was observed when one or more hydration layers are present around the enzyme.
8. The chemical synthesis of both enantiomers of two pseudo-diglycerides and six pseudo-triglycerides have been performed and these substrate analogues have been used to characterize the positional and stereopreference of several lipases purified by BRIDGE's partners.
9. Potential applications of the oil-drop tensiometer have been developped for studying the kinetics of (phospho)lipase action.
10. Epitope mapping and immunoinactivation of human gastric lipase were achieved using bulk and monomolecular film techniques.

MAJOR COOPERATIVE LINKS

While the project has been running it can be observed that collaboration among the project teams has been increasing and that the interactions among researchers in Europe have become more frequent due to the existence of ou stimulating BRIDGE T-Lipase project.

Exchanges of staff and students:

— *Long term (>1 year):*
Ransac, S. (Marseille → Groningen);
Martinez, C. (Marseille → Vlaardingen, Utrecht);
Carrière, F. (Marseille → Copenhagen);
Longhi, S. (Milano → Marseille).

— *Short term (<1 year):*
Ladefoged, C. (Copenhagen→ Marseille);
Melo, E. (Lisbon → Marseille);
Ransac, S. (Groningen → Bochum).

The closing meeting of the BRIDGE T-Lipase programme took place in a beautiful setting at the Bendor Island (France) from September 14-17th 1994. It attracted about 180 participants, which from many perspectives is a respectable number, indicating a high level of interest in the topic. Fifty one of the participants had an industrial affiliation. The fact that many young scientists participated confirmed that many laboratories, both within industry and academia regard the challenging and complex problems associated with lipase structure and function as a very timery and exciting area of research.

The meeting confirmed in a very convincing way that the concerted scientifid efforts that have taken place in the EC BRIDGE T-Lipase programme have led to a very dramatic progress in out understanding of the structure and function of these complex enzymes.

To our knowledge no comparable effort has taken place in other regions of the world. Several of the US participants were impressed with the size of the effort of the BRIDGE T-Lipase programme as well as by the quality of the presentations. Some commented that such an effort could not take place in the US.

PUBLICATIONS

Joint publications

Cudrey C., van Tilbeurgh H., Gargouri Y. and Verger R. (1993). Inactivation of pancreatic lipases by amphiphilic reagents (C12-TNB, THL). Dependence upon partitionning between micellar and oil phases. *Biochemistry,* **32**: 13800-13808.

Martinez C., Nicolas A., van Tilbeurgh H., Egloff M.P., Cudrey C., Verger R. and Cambillau C. (1994). Cutinase, a lipolytic enzyme with a preformed oxyanion hole. *Biochemistry,* **33**: 83-89.

Thirstrup K., Verger R. and Carrière F. (1994). Evidence for a pancreatic lipase subfamily with new kinetic properties. *Biochemistry,* **33**: 2748-2756.

Carrière F. Thirstrup K., Boel E., Verger R. and Thim L. (1994). Structure-function relationships in naturally occuring mutants of pancreatic lipase. *Protein Engineering,* **7**: 563-569.

Egloff M.P., Marguet F., Buono G., Verger R., Cambillau C. and van Tilbeurgh H. The 2.4Å resolution structure of the pancreatic lipase colipase complex inhibited by a C_{11}-alkyl phosphonate. *Biochemistry* (1994). (in press).

Jaeger K.-E., Ransac S., Koch H.B., Ferrato F. and Dijkstra B.W. (1993). Topological characterization and modelling of the 3D structure of lipase from *Pseudomomas aeruginosa. FEBS Lett.,* **332**: 143-149.

De Haas G.H., Dijkman R., Lugtigheid R.B., Dekker N., Van der Berg L., Egmond M.R. and Verheij H.M. (1993). Competitive inhibition of lipolytic enzymes. IX. A comparative study on the inhibition of pancreatic phospholipases A_2 from different sources by (R)-2-acylamino phospholipid analogues. *Biochim. Biophys. Acta,* **167**: 281-288.

Leuveling Tjeenk M., Bulsink Y.B.M., Slotboom A.J., Verheij H.M., De Haas G.H., Demleitner G. and Götz F. (1994). Inactivation of *Staphylococcus hyicus* lipase by hexadecylsulfonyl fluoride: evidence for an active site serine. *Protein Engineering,* **7**: 579-583.

MeloE.P., Ivanova M.G., Aires-Barros M.R., Cabral J.M.S. and Verger R. 'Glyceride syntheis catalysed by cutinase using the monomolecular film technique'. *Biochemistry* (submitted).

Taipa M.A., Liebeton K., Costa J.V., Cabral J.M.S. and Jaeger K.E. 'lipase from *Chromobacterium viscosum*: Biochemical Characterization Indicating Homology to Lipase from *Pseudomonas glumae'. Biochim. Biophys. Acta* (submitted).

Breg J.N., Sarda L., Cozzone P.J., Boelens R., Rugani N. and Kaptein R. Solution structure of porcine pancreatic procolipase as determined from ^1H homonuclear 2D- and 3D-NMR experiments at 500- and 600-MHz *Eur. J. Biochem.* (in press) 1994.

Rugani N., Carrière F., Thim L., Borgström B. and Sarda L. Lipid binding and activating properties of porcine pancreatic colipase split at the Ile_{79}-Thr_80 bond *Biochim. Biophys. Acta.* (in press) 1994.

Individual publications

Ivanova Tz., Raneva V., Panaiotov I. and Verger R. (1993). Kinetics of spreading of DOPC liposomes at the air-water interface subjected to phospholipase A_2 hydrolysis. *Colloid Polym. Sci.,* **271**: 290-297.

Aoubala M., Douchet, I., Laugier R., Hirn M., Verger R. and de Caro A. (1993). Purification of human gastric lipase by immuno affinity and quantification of this enzyme in the duodenal contents using a new ELISA procedure. *Biochim. Biophys. Acta,* **1169**: 183-188.

Gargouri Y., Bensalah A., Douchet I. and Verger R. (1994). Kinetic behaviour of five pancreatic lipases using emulsions and monomolecular films of synthetic glycerides. *Biochim. Biophys. Acta.* (in press).

Marguet F., Cudrey C., Verger R. and Buono G. (1993). Digestive lipases: inactivation by phosphonates. *Biochim. Biophys. Acta,* **1210**: 157-166.
Laurent S., Ivanova M., Pioch D., Graille J. and Verger R. (1994). Interactions between β-cyclodextrin and insoluble glyceride monomolecular films at the argon/water interface. Application to lipase kinetics. *Chem. Phys. Lipids.* **70**: 35-42.
Carrière F., Gargouri Y., Moreau H., Ransac S., Rogalska E. and Verger R. (1994). Gastric lipases: Cellular Biochemical and Kinetic Aspects. In: *Lipases: their structure, biochemistry and application.* Wooley P. and Petersen S.B. eds. Cambridge University Press, Cambridge (pp. 182-205).
de la Fournière L., Ivanova M.G., Blond J.P., Carrière F. and Verger R. (1994). Surface behaviour of human pancreatic and gastric lipases. *Colloids ans Surfaces B: Biointerfaces.* **2**: 585-593.
Labourdenne S., Gaudry-Roland N., Letellier S., Lin M., Cagna A., Esposito G., Verger R. and Rivière C. (1994). The oil-drop tensiometer: potential applications for studying the kinetics of (phospho)lipase action. *Chem. Phys. Lipids.* **7**: 163-173.
Aoubala M., Bonicel J., Bénicourt C., Verger R. and de Caro A. (1994). Tryptic cleavage of gastric lipases. Location of the single disulfide bridge. *Biochim. Biophys. Acta.* **1213**: 319-324.
Oliveira A.C. and Cabral J.M.S. (1993). 'Kinetic Studies of *Mucor miehei* Lipase in Phosphatidylcholine Microemulsions'. *J. Technol. Biotechnol.* **56**: 247-252.
Taipa M.A., Aires-Barros M.R. and Cabral J.M.S. (1992). 'Purification of Lipases from *Chromobacterium viscosum*: Is there a High Molecular Weight Lipase?' *Biotechnol. Tech.* **6**: 533-538.
Prazeres D.M.F., Garcia F.A.P. and Cabral J.M.S. (1993). 'An Ultrafiltration Membrane Bioreactor for the Liposys of Olive Oil in a Reversed Micellar Media'. *Biotechnol. Bioeng.* **41**: 761-770.
Prazeres D.M.F., Garcia F.A.P. and Cabral J.M.S. (1993). 'Colorimetric Assays for the Determination of Dioctyl Sodium Sulphosuccinate and Glycerol in Reversed Micelles'. *Biotechnol. Tech.* **7**: 14.
Prazeres D.M.F., Garcia F.A.P. and Cabral J.M.S. (1993). 'Modeling Lipolysis in Reversed Micellar System: Part I — Conventional Batch Reactor'. *Biotechnol. Bioeng.* **42**: 759-764.
Prazeres D.M.F., Garcia F.A.P. and Cabral J.M.S. (1993). 'Modeling Lipolysis in Reversed-Micellar System: Part II — Membrane Reactor'. *Biotechnol. Bioeng.* **42**: 765-771.
Sebastiao M.J., Cabral J.M.S. and Aires-Barros M.R. (1993). *'Fusarium solani pisi* Recombinant Cutinase Partitioning in PEG/potassium Phosphate Aqueous Two-phase Systems', *Biotechnol. Tech.* **7**: 631-634.
Taipa M.A., Moura-Pinto P. and Cabral J.M.S. (1994). 'Identification and Characterization of *Chromobacterium viscosum* Lipase', *Biotechnol. Tech.* **8**: 27-32.
Carneiro-da-Cunha M.G., Aires-Barros M.R., Tambourgi E.B. and Cabral J.M.S. (1994). 'Recovery of a Recombinant Cutinase with Reversed Micelles in Continuous Perforated Rotating Disc Contactor', *Biotechnol. Tech.* **8**: 413-418.
Sebastiao M.J., Cabral J.M.S. and Aires-Barros M.R. (1994) 'Partitioning of Recombinant *Fusarium solani pisi* Cutinase in Polyethylene Glycol-Aqueous salt Two-Phase Systems', *J. Chromat.* **668**: 139-144.
Pinto-Sousa A.M.C., Cabral J.M.S. and Aires-Barros M.R. 'Ester Synthesis by a Recombinant Cutinase in Reversed Micelles of a Natural Phospholipid, *Biocatalysis.* **9** 1884) 169-179.
Carneiro-da-Cunha M.G., Cabral J.M.S. and Aires-Barros M.R. 'studies on the extraction and Back-extraction of a Recombinant Cutinase in a Reversed Micelles Extraction Process', *Bioprocess Eng.* **11** (1994) 203-208.
Melo E.P., Aires-Barros M.R. and Cabral J.M.S. 'Triglycerides Hydrolysis and Stability of Recombinant Cutinase from *Fusarium solani* in AOT-Isooctane Reversed Micelles'. *Appl. Biochem. Bioctechnol.* (in press).
Dijkman R., Cox R., Van den Berg L., Verheij H.M. and De Haas G.H. (1994). competitive inhibition of lipolytic enzymes. X. Further delineation of the active site of pancreatic phospholipases A_2 from pig, ox and horse by comparing the inhibitory power of a number of (R)-2-acylamino phospholipid analogues. *Biochim. Biophys. Acta.* **1212**: 50-58.
Dezan C., Daniel C., Hirn J., Sarda L. and Bellon B. Production of monoclonal antibodies to human pancreatic procolipase and characterization of antibodies by competitive binding studies. Identification of a neutralizing antibody which recognizes a conformational epitope related to colipase activity *Hybridoma*, **13** (1994) 509-517.

Cambilliau C. and Van Tilbeurgh H. (1993). Structure of hydrolases: lipases and cellulases. *Current Opinion in Structural Biology.* 3: 885-895.
Van Tilbeurgh H., Roussel A., Lalouel J.M. and Cambilliau C. (1994). Lipoprotein lipase. Molecular model based on the pancreatic lipase X-ray structure: consequences for heparin binding and catalysis. *J. Biol. Chem.* **269**: 4626-4633.
Charlemagne D. and Legoy M.D. Enzymatic synthesis of polyglycerol-fatty acid esters in a solvent-free system. *J. Am. Oil Chem. Soc.* (1995). (in press).
Drouet P., Thomas C. and Legoy M.D. Production of 13(S)-Hydroperoxy-9(Z),11(E)-octadecadienoic acid using soybean lipoxygenase 1 in a biphasic octane-water system. *Tetrahedron Let.* **35** (1994) 3923-3926.
Lamare S. and Legoy M.C. Working at controlled water activity in a continuous process: the bas/solid system as a solution. *Biotechnol. Bioeng.* 1995. (in press).

Joint patent publications

Hjorth A. Wöldike H.F., Thim L., Gormsen E. and Verger R. Mammalian pancreatic lipase and variant thereof. Patent No. WO 93/00426.
Verger R., Rivière C., Esposito G. and Cagna A. Procédé de mesure de l'activité des phospolipases. Brevet I.N.P.I. N° 9211474.

Molecular structure and specificity relationship of microbial triacylglycerol lipases — 'pre-final report' (BIOT CT-910258)

COORDINATOR:
R.D. SCHMID, Institut für Technische Biochemie, Univ. Stuttgart, Stuttgart, D
PARTICIPANTS:
C. DAVIES, Unilever Research, Colworth House, GB
F. SPENER, Universität Münster, Münster, D
J. CABRAL, Instituto Superior Tecnico, Lisboa, P
L. ALBERGHINA, Univ. Studi di Milano, Milano, I
E. CERNIA,Univ. Studi di Roma, Roma, I
R. VERGER, CNRS, Marseille, F
F. PALTAUF, Technische Universität Graz, Graz, AT

BACKGROUND INFORMATION

The present project has been concerned with the structure and mechanism of various microbial lipases. Microbial lipases have found wide applications in industry, e.g. in household detergents and in the synthesis of specialty chemicals, and it is inferred that a better understanding of their function will lead to novel industrial processes.

OBJECTIVES AND PRIMARY APPROACHES

To clone, express and sequence microbial lipases, to investigate their molecular structure and to study their catalytic mechanism. To elucidate their differences in specificity. To find clues for improving their function, via protein engineering, for industrial applications.

RESULTS AND DISCUSSION

1. Introduction

The scope of this project was the elucidation of structure, catalytic mechanism and enzymatic properties of selected microbial lipases. The lipases investigated were isolated from *Rhizopus oryzae* (ROL), *Candida rugosa* (formerly *Candida cylindracea)* (CRL), *Geotrichum candidum* (GCL), *Chromobacterium viscosum* (CVL), *Pseudomonas spec* (PL) and *Bacillus thermocatenulatus* (BTL).

2. Purification, crystallization and characterization

Table 1
Purification and crystallization of lipases

Lipase from	*Abbreviation*	*Spec. Activity of pure lipase (U/mg protein)*	*Crystals, remarks*	*Laboratory*
Rhizopus oryzae	ROL		not successful	Münster, GBF
Candida rugosa	CCL		not attempted	Roma
Geotrichum candidum A, C	GCL-A, GCL-C		diffracting crystals	Münster, GBF
Geotrichum candidum B	GCL-B		small diffracting crystals of poor quality	GBF, Unilever
Chromobacterium viscosum	CVL		diffracting crystals, 4 isomorphous heavy metal derivatives	Münster, GBF
Pseudomonas cepacia	PCL		diffracting crystals	GBF
Bacillus thermocatenulatus	BTL		not (yet) attempted	GBF

As purification and crystallization of the lipases were crucial steps for ensuing x-ray analysis and cloning, most participating laboratories were initially involved at this endeavour. The division of labor and the results are summarized in Table 1.

3. Cloning and expression

The results obtained by our project group are summarized in Table 2.

Table 2

Lipase from	*Code*	*DNA sequence*	*Expression system*	*Remarks*	*Laboratory*
Rhizopus oryzae	ROL	yes	E. coli	Inclusion body	GBF
Candida rugosa	CRL	yes	S. cerevisiae	unusual codon usage	MILANO
Geotrichum candidum	GCL-B	yes	S. cerevisiae		UNILEVER
Bacillus thermocatenulatus	BTL-B	yes	E. coli	under study	GBF

Rhizopus oryzae lipase (ROL): the gene was cloned and overexpressed in *E. coli* at the GBF from the genomic DNA of *Rhizopus oryzae* DSM853.
Candida rugosa lipase (CRL): the aim of this project, carried out by MILANO, was to clone and characterize the lipase gene(s) from *C. rugosa* and to overexpress the cloned proteins in *Saccharomyces cerevisiae. C. rugosa* shows a peculiar use of the genetic code, in that the universal codon for leucine CUG is used for serine. As a prerequisite for any gene manipulation, extensive mutagenesis with the purpose to obtain functional enzymes upon expression in host organisms was carried out.

Geotrichum candidum lipase (GCL-B): at the UNILEVER laboratories, GCL-A and GCL-B clones were isolated and expressed in *E. coli* and *S. cerevisiae.*
Bacillus thermocatenulatus lipase: at the GBF, a 43 kDa thermophilic lipase of *Bacillus thermocatenulatus* DNA cloned and expressed in *E. coli.* The enzyme was analyzed as to its sequence and its enzymatic properties

4. Structure elucidation and modelling

The results obtained by our project group are summarized in Table 3.

Rhizomucor miehei lipase (RML): RML structure could be solved by use of the Cα-coordinates from the Brookhaven data base (code name 1TGL), and a preliminary model of the active ('open lid') form of RML was constructed. Docking of trilaurylglycerol led to a first model of a RML/ substrate complex.
Rhizopus oryzae lipase (ROL): starting from the x-ray structure of *Rhizomucor miehei* lipase (RML), a structural model of ROL (55% sequence identity with RML) was constructed at the GBF by homology modeling. Several variants were prepared by site-directed mutagenesis, and information on the catalytic mechanism of the enzyme was obtained.

Candida rugosa lipase (CRL): A model of the LIP1 protein was derived by homology from the structure of the *G. candidum* enzyme. Lipase isoforms subsequently purified from a commercial preparation by ROMA showed biochemical properties and sequences of their amino terminus consistent with those predicted. Differences in the catalytic behaviour of lipase species were also reported. As *Candida rugosa* uses the codon CUG for serine instead of leucine and in this respect deviates from universal codon usage, mutagenesis of this codon is a prerequisite to expressing the CRL genes in standard host organisms. In order to mutagenize the 19 CUG codons occuring in CRL in a more rational way, a 3D model of CRL was used to identify those leucine residues important for the enzyme structure, thus restricting the first set of codons to be mutated to 9. Mutagenesis/expression experiments are now under way.

Geotrichum candidum lipase (GCL-A and GCL-B): by cooperation of UNILEVER and GBF, a model of GCL-B was constructed , using the published coordinates of GCL ATCC 34614.

Chromobacterium viscosum lipase (CVL): CVL structure was resolved at the GBF at 1.6 Å resolution. A complex of CVL with the lipase inhibitor lipstatin, kindly offered by Dr. Winkler (Hoffmann LaRoche), was prepared by MÜNSTER and crystallized by GBF, but did not diffract.

Table 3

Lipase from	*Code*	*Structure-related studies*	*Results*	*Laboratory*
Rhizomucor miehei	RML	x-ray analysis based on published data for the Cα trace	crystal structure at Å resolution	GBF
Rhizopus oryzae	ROL	model based on homology to Rhizomucor miehei	model established and corroborated by site-directed mutagenesis	GBF
Candida rugosa	CRL	models based on x-ray coordinates of *Geotrichum candidum* and *Candida rugosa*	model established and utilized for site-directed mutagenesis of serine sites	MILANO
Geotrichum candidum	GCL-C	model based on x-ray data-set of native enzyme and the coordinates of *Geotrichum candidum* lipase ATCC 34614	in progress	GBF
	GCL-B	model based on homologous replacement of *Geotrichum candidum* lipase ATCC 34614		UNILEVER, GBF
Chromobacterium viscosum	CVL	x-ray structure determination	structure resolved at 1.6 Å	GBF
Pseudomonas cepacia	PCL	model based on x-ray data-set of native enzyme and homologous replacment of Chromobacterium viscosum lipase	in progress	GBF

5. Properties, mechanism of action and protein engineering

The results obtained by our project group are summarized in Table 4.

Table 4

Lipase from	*Code*	*Type of study*	*Laboratory*
Rhizopus arrhizus, Rhizopus oryzae	RAL, ROL	kinetic properties, stereopreference, inhibitors, protein engineering	Graz, Münster, Marseille, GBF
Candida rugosa	CRL	kinetic properties, stereopreference, inhibitors	Graz, Münster, Marseille, Roma
Geotrichum candidum	GCL	kinetic properties, substrate specificity	UNILEVER, Marseille
Chromobacterium viscosum	CVL	kinetic properties, substrate specificity	Graz, Münster, Marseille, Lisboa
Pseudomonas spec.	PCL	kinetic properties, substrate specificity	Graz, Münster, Marseille
Bacillus thermocatenulatus	BTL	stability, substrate specificity	GBF

MAJOR SCIENTIFIC BREAKTHROUGHS

1. One lipase structure solved (CVL) (under progress: PCL and GCL-C), several lipase structures modelled (ROL, CRL, GCL-B),
2. Several lipases cloned and sequenced (ROL, CRL, GCL-B, BTL),
3. Lipase mechanism elucidated by site-directed mutagenesis (ROL, CRL),
4. CRL mutants prepared for expression in standard host,
5. Substrate specificity of ROL investigated by molecular mechanics calculations,
6. Lipase mechanism and stereopreference addressed by kinetic studies (RAL, CRL, CVL, PCL)

MAJOR COOPERATIVE LINKS

Meetings of the project group in Braunschweig (1990), Braunschweig/Münster (1991), Capri (1992) and Elsinore (1993). Various meetings of subgroups.

1. MILANO: with ROMA for purification and activity assays of the mutated proteins, with MARSEILLE (Prof. Cambillau) for the exchange of researchers and the x-ray analysis of lipase mutants,
2. LISBOA: with MÜNSTER for the exchange of a Ph. D. student,
3. UNILEVER: individual visits to MARSEILLE, MILANO. Samples to GBF, MÜNSTER, MARSEILLE, GRAZ and MILANO,
4. MÜNSTER: exchange of researchers with GRAZ, GBF, LISBOA.
5. GBF: individual visits to MARSEILLE, MILANO, ROMA, GRAZ. Exchange of samples with MÜNSTER, GRAZ, MARSEILLE, UNILEVER.

PUBLICATIONS

Joint publications

Hedrich, H.C., Spener, F., Menge, U., Hecht, H.-J., Schomburg, D. and Schmid, R.D, *Enzyme Microb. Technol.* (1991), **13**, 840-847

Hedrich, H.C., Isobe, K., Stahl, B., Nokihara, K., Kordel, M., Schmid, R.D., Karas, M., Hillenkamp, F. and Spener, F. (1993), Matrix-assisted ultraviolet laser desorption/ionization mass spectrometry applied to multiple forms of lipase. *Anal. Biochem.* **211**, 288-292

Lang. D., Hofmann, B., Haalck, L., Hecht, H.-J., Spener, F., Schmid, R.D. and Schomburg, D. (1994), Crystallization and Preliminary X-ray Analysis of a Lipase from *Chromobacterium viscosum*. *Acta Cryst.* **50**, 272-274

Stadler, P., Kovac, A., Haalck, L., Spener, F. and Paltauf, F. (1994) Stereoselectivity of microbial lipases (submitted).

Cudrey, C., Ferrato, F., Haalck, L., Paltauf, F., Rogalska, E., Spener, F., Stadler, P. and Verger, R. (1994) Stereoselective hydrolysis of triglycerides and triglyceride-analogs by animal and microbial lipases. Proc. 6th Eur. Congr. Biotechnol., L. Alberghina, L. Frontali, P. Sensi (eds.), pp. 137-140

Alberghina L., Schmid D.R., Verger R. Edts, GBF monograph vol. 16 pp. 369-372 (1991), '*Candida cylindracea* lipase: approaches towards the understanding of structure-performance relationships', Alberghina L., Battinelli L., Cernia E., Fusetti F., Lotti M., Soro S. ECB6 Proceeding of the 6th European Congress on Biotechnology Elsiver 129-132 (1994)

Individual publications

Haalck, L., Hedrich, H.C., Hassink, J. and Spener, F. (1992) Modification of waste fats by lipase-catalyzed reaction in solvent-free substrate blends. *Progr. Biotechnol.* **8**, 505-512

Cernia, F. Gasparrini, D.Misiti, C. Palocci, 'Enantioselectivity and reactivity of immobilized lipase in supercritical carbon dioxide' *J. Molecular Catalysis*, vol.89, issue No. 1-2, pp. 11-18, April 1994

Cernia, M. Delfini, A. Magri', C. Palocci, 'Enzymatic catalysis by lipase from *Candida cylindracea*: enantiomeric activity evaluation by ^{1}H- and ^{13}C-NMR.' *Cellular and Molecular Biology* **40 (2)**, 193-199,1994

Cernia, F. Gasparrini, D. Misiti, C. Palocci, *'Pseudomonas sp.* lipase catalytic activity in supercritical carbon dioxide'. *Chem. Biochem. Eng.* Q. **8(1)**, 1-4, 1994

Charton, E., Davies, C. and Macrae, A.R. (1992) Use of Specific Polyclonal Antibodies to Detect Heterogeneous lipases from *Geotrichum candidum*. *Biochem. et Biophys. Acta.* **1127**, 191-198.

Charton, E., Macrae, A.R. (1993) Specificities of Immobilised *Geotrichum candidum* CMICC 335426 lipases A and B in hydrolysis and ester synthesis reactions in organic solvents. *Enzyme Microb. Technol.* **15**, 489-493.

Taipa, M.A., Aires-Barros, M.R. and Cabral, J.M.S. (1992) *Biotechnol. Tech.* **6**, 533-538.

Taipa, M.A., Moura-Pinto, P. and J.M.S.Cabral (1994a) *Biotechnol. Tech.* **8**, 27-32.

Taipa, M.A., Liebeton, K., Costa, J.V., Cabral, J.M.S. and Jaeger, K.E. (1994b), (submitted).

Longhi S., Fusetti F., Grandori R., Lotti M., Vanoni M., Alberghina L. (1992) 'Cloning and nucleotide sequences of two lipase genes from *C. cylindracea'. Biochim. Biophys. Acta* **1131**: 227-232

Longhi S., Lotti M., Fusetti F., Pizzi E., Tramontano A., Alberghina L. (1992) 'Homology-derived three-dimensional structure prediction of *C. cylindracea* lipase'. *Biochim. Biophys. Acta* **1165**: 129-133

Lotti M., Grandori R., Fusetti F., Longhi S., Brocca S., Tramontano A., Alberghina L. (1993) 'Cloning and analysis of *C. cylindracea* lipase sequences'. *Gene* **124**: 45-55

Lotti M., Tramontano A., Longhi S., Fusetti F., Brocca S., Pizzi E., Alberghina L. (1994) 'Variability within the C. rugosa lipases family'. *Prot. Eng.* 7: 531-535

Haalck L. and Spener F. (1992) Application of FPLC in the separation of industrial lipases. In 'Downstream Processing in Biotechnology', R.N. Mukherjea (ed.), Tata-McGraw Hill Publ., Bombay, pp. 33-42

Cernia, A.D. Magrì, G. Ortaggi, M. Castagnola, R. Rabino, F. Bartoli, 'Lipolytic enzymes separation and purification through functionalized syntetic polymers'. Lipase: Structure, mechanism and genetic engineering (in press).

Schmidt-Dannert, H. Sztajer, W. Stöcklein, U. Menge, R.D. Schmid; Screening, purification and properties of a thermophilic lipase from *Bacillus thermocatenulatus*; 1994, *Biochim. Biophys. Acta* (in press).
A. Borioni, E. Cernia, A.D. Magrì, C. Palocci, S. Soro, 'Some influence of the chemical composition on the biocatalyzed hydrolysis of organic esters', Proceedings of the Workshop: Lipases: structure, function and applications on biotransformations, Coventry, July 16-18, 1991
Cernia E., Battinelli L., Soro S, 'Monolayer Technique for studing lipase activity', Proteine '92 Pavia Maggio 1992
Cernia E., Castagnola M., Ortaggi G., Bersani M. T., Rabino R., Battinelli L., Soro S., 'Affinity chromatography and other purification methods for lipolytical enzymes: a comparison', ECB6 Firenze June 1993
Battinelli L., Cernia E., Soro S., 'Hydrolysis of triacylglycerol monolayers biocatalyzed by Candida cylindracea lipase', LB6 Trois-Rivieres, Quebec, Canada July 1993
Botta M., Cernia E., Corelli F., Manetti F., Soro S., 'Computational studies on the mechanism of 2-arylpropionic esters hydrolysis catalized by Candida cylindracea lipase', Lipases: Structure, function and protein engineering. Elsinor, Denmark October 1993
Cernia E., Ortaggi G., Battinelli L., Bersani M.T., , Soro S., 'The esters of cross-linked polyvinyl alcohol as selective media for lipases separations', Protein separations in polymeric media, San Diego, March 1994Alberghina L. and Lotti M. (1992) 'Yeast: from genetic engineering to protein engineering'., Biotec 6/92: 36-40
Fusetti F., Lotti M., Brocca S., Longhi S., Alberghina L. (1992) 'Structure of a family of lipase-encoding genes from C. cylindracea'., Med Fac. Landbouww. Univ. Gent 57/4b: 2045-2052
Alberghina L., Battinelli L., Cernia E., Fusetti F., Longhi S., Lotti M., Soro S. (1994) '*C. cylindracea* lipase: approache towards the understanding of structure-performance relationships'., ECB6: Proceedings of the 6th European Congress on Biotechnology, Elsevier: 129-132

T-PROJECT

'LACTIC ACID BACTERIA'

Improvement and exploitation of lactic acid bacteria for biotechnology purposes (BIOT CT-910263)

COORDINATOR:
C. DALY, NFBC, University College, Cork, IRL
PARTICIPANTS:
C.Y. BOQUIEN, INRA, Thiverval Grignon, F
V. BOTTAZZI, Univ. Cattolica del Sacro Cuore, Piacenza, I
M.C. CHOPIN, INRA, Jouy-en-Josas, F
T.M. COGAN, Teagasc, Fermoy, Cork, IRL
M.A. COLLINS, Queens University of Belfast, Belfast, GB
P.E.HERNÁNDEZ CRUZ, Universidade Complutense, Madrid, E
C. DALY/G.F. FITZGERALD, University College, Cork, IRL
R. JIMENEZ DIAZ, U.E. de I. de Biotecnologia de Alimentos, Sevilla, E
A. MENDES FAIA, Universidade de Trás-os-Montes e Alto Douro, Vila Real, P
M.J. GASSON, AFRC Institute of Food Research, Norwich, GB
K.J. HELLER, Bundesanstalt für Milchforschung, Kiel, D
W.P. HAMMES, Universität Hohenheim, Stuttgart, D
E.B. HANSEN, Chr. Hansen's Laboratorium, Hørsholm, DK
B. JELLE, Chr.Hansen's Laboratorium, Hørsholm, DK
J. JOSEPHSEN, Royal Veterinary & Agric. Univ., Frederiksberg, DK
A. JOYEAUX, Société Equipharm, Saint-Simon, F
G. KALANTZOPOULOS, Agric. Univ. of Athens, Athens, GR
K. KERSTERS, Universiteit Ghent, Ghent, B
W.N. KONINGS, University of Groningen, Haren, NL
A.M. LEDEBOER, Unilever Research Lab., Vlaardingen, NL
V.M. MARSHALL, University of Huddersfield, Huddersfield, GB
F. MULHOLLAND, AFRC Inst. of Food Research, Reading, GB
I.F. NES, Agricultural University, Ås, N
R. PLAPP, Universität Kaiserslautern, Kaiserslautern, D
J.R. PRIGENT, Texel (Rhône-Poulenc Group), Dangé-St-Romain, F
P. RITZENTHALER, LMGM du CNRS, Toulouse, F
G. SAKELLARIS, Nati. Hellenic Res. Foundation, Athens, GR
M.H. SANTOS, Centro Técnol. Quimica e Biolog., Oeiras, P
K.H. SCHLEIFER, Technische Universität München, München, D
J.E. SUAREZ, Universidad de Oviedo, Oviedo, E
J.H.J. HUIS IN'T VELD, TNO Nutrition & Food Res., Zeist, NL
G. VENEMA, University of Groningen, Haren, NL
W.M. DE VOS, NIZO, Ede, NL
P.J. WARNER, Cranfield Institute of Technology, Bedford, GB

BACKGROUND INFORMATION

This project involved a transnational collaboration of 34 research groups dedicated to studying the Lactic Acid Bacteria (LAB) at a fundamental and applied level. These bacteria play an essential role in a range of dairy, meat and vegetable fermentations which are of major economic importance to the European economy. The scientific effort within the T-project was directed to addressing a number of key areas regarding food cultures which were intended to relieve bottlenecks in production and process applications, enhance culture performance and efficiency

in industrial situations and generate cultures with novel characteristics for new applications. Thus, a number of groups constituting the CORE used a multidisciplinary approach, including physiology, biochemistry and molecular biology to advance the fundamental knowledge of key culture activities required for successful food fermentations. These and other groups were assembled into four LEAFS each of which targeted key industrially important traits: proteolysis; bacteriophage and bacteriophage resistance; antimicrobials; and metabolism/screening. A major strength of the total project was the cross-fertilization between all participant groups which maximized the yield in terms of scientific output and knowledge of the LAB.

A. CORE TECHNOLOGIES

BACKGROUND INFORMATION

The CORE activities in the project were designed to underpin the LEAF activities which have a more direct focus on application. At the start of the project there was a need to further develop genetic technology and this was the main theme of the CORE. Particular emphasis was placed on the control of gene expression of biotechnologically important traits of the lactic acid bacteria.

OBJECTIVES AND PRIMARY APPROACHES

1. Characterization of homologous gene expression in lactic acid bacteria and its exploitation for the expression of heterologous genes.
2. Development of technologies for the chromosomal integration of DNA.
3. Characterization and exploitation of plasmid replication.
4. Characterization of gene transfer systems with an emphasis on conjugation.
5. Isolation and analysis of new genes from lactic acid bacteria.

RESULTS AND DISCUSSION

1. Gene expression

Understanding the control of gene expression has advanced in several ways. Improved systems to monitor gene expression were developed by using the luciferase genes (*luxAB*) from *Vibrio fischeri* and the β-glucuronidase gene (*gusA*) from *E. coli*. The value of the new systems was established in analyses of expression and control in the lactococcal lactose operon and proteinase genes, respectively. In addition, the staphylococcal nuclease gene was used to monitor secretion.

Lactose operon regulation was characterised in detail, building on the discovery of a divergently transcribed repressor gene *lacR*. Two operator sites were defined in footprinting studies and protein engineering of the LacR protein was used to locate the recognition helix and to generate mutants with improved operator binding. The inducer of the lactose operon was shown to be tagatose-6-phosphate, and LacR protein engineering revealed the charged residues involved in its response to the inducer. Catabolite repression was shown to be an additional aspect of lactose operon regulation and a glucose responsive element was defined.

The lactose operon promoter was exploited in the development of plasmid expression vectors which have been used for the controlled expression of a variety of different genes in *L. lactis* including a novel lysin gene that was potently active against *Listeria monocytogenes*. A strain of *L. lactis* with a chromosomally integrated copy of the lactose operon was exploited to optimize food-grade plasmid vector systems

and to investigate the controlled expression of integrated target genes. Stable amplification of heterologous genes was also achieved.

The control of gene expression was investigated in several different systems. Lactococcal genes for the biosynthesis of amino acids were characterised in substantial detail and the mechanisms involved in the control of the *trp, his* and *leu-ilv* operons were revealed. Classical control of transcription by repressors as well as the first examples in the lactic acid bacteria of transcription attenuation were described. In addition, a new type of regulatory mechanism involving a metabolic shunt in the branched chain amino acid biosynthetic pathway was discovered. This involved the *aldB* gene for acetolactate decarboxylase which together with the *ilvBN* genes was also exploited in the Metabolism LEAF for metabolic engineering of the diacetyl pathway. A gene *bglR*, encoding a regulatory protein of the family of RNA-binding transcription antiterminators, was also characterised in the *trp* operon. The DNA sequence information gained in this study of amino acid biosynthetic operons was exploited for a molecular taxonomic study of the *lactis* and *cremoris* subspecies of *L. lactis* thereby revealing frequent misclassification of strains. Notably this included the key genetic strain *L. lactis* 712. In addition, the inability of dairy strains of *L. lactis* to synthesize histidine or isoleucine, leucine and valine was explained by the accumulation of mutations in the relevant operons.

Genetic switches are of particular relevance for the future application of genetic engineering technology. DNA sequence analysis of the complete genome of a temperate lactococcal bacteriophage led to the characterization of the regulatory system controlling the balance between the lysogenic and lytic states. A repressor gene and a gene encoding a protein homologous to the bacteriophage lambda Cro protein were found. Gene fusions to *lacZ* were used to demonstrate that this region provided a clean genetic switch and its potential was established by demonstrating controlled expression of the bacteriophage lysin and holin genes. Environmental control of gene expression was established by cloning and characterising the heat shock gene analogous to *dnaK* and the use of integrative gene fusions to identify a salt regulated promoter. The *recA* gene of *L. lactis* was studied leading to the interesting observation that the RecA protein plays a role in the repair of oxygen radical induced DNA damage.

2. Chromosomal integration

Chromosomal integration is of importance both as a strategy for the stable food-grade cloning of genes and because a range of genetic elements naturally interact with the chromosome. With respect to the latter, pulsed field gel electrophoresis (PFGE) of *L. lactis* chromosomal DNA was used to map sites for the integration of insertion sequences, nisin transposons, a fragment of the metabolic plasmid pLP712, a sex factor involved in high frequency conjugation and a temperate bacteriophage. PFGE was also used to locate copies of transposons Tn916 E and Tn1545 that were integrated within the chromosome of enteric species of *Lactobacillus*.

A sophisticated and versatile system for the integration of cloned genes into the chromosome was developed based on conditional replicating derivatives of the lactococcal plasmid pWV01. This relied upon characterization of the replication mechanism of this plasmid and involved the separation of the trans acting control gene *repA* from its target. Using this system a range of genetic technologies were perfected, including efficient food-grade stabilisation of genes, delivery of defined

mutations by replacement recombination, random mutagenesis by gene disruption and the construction of transcriptional fusions to reporter genes.

3. Plasmid replication

Analysis of DNA replication has been continued with new molecular characterization of the replicons of lactococcal plasmids pCI305, pCI528 and pWV02 and the *Leuconostoc lactis* plasmid pCI412. Whilst the latter *Leuconostoc* plasmid appeared to possess another rolling circle replicon the three lactococcal plasmids replicated via a theta mechanism. These replicons are of special importance in offering enhanced stability in plasmid cloning vectors, especially where derivatives carry long inserts. This was experimentally demonstrated to be true for pWV02 which was then used to build a new food-grade cloning and expression vector.

4. Gene transfer — conjugation

Natural gene transfer processes are of particular importance both as an alternative strain development technology to genetic engineering and because they are of relevance to the genetic containment of Genetically Engineered Microorganisms. The high frequency conjugation system of *L. lactis* 712 was subjected to in depth molecular analysis with the important discovery that it was controlled by a chromosomally located sex factor. In *Lactobacillus* another high frequency conjugation system, which also characteristically involved a cell aggregation phenomenon, was characterised. In this case a secreted protein named 'Aggregation Promotion Factor (APF)' was shown to recognize and bind to receptors located on the outermost envelopes of bacterial cells. Polyclonal antibodies to APF were used to reveal the existence of immunologically related APF's in further enteric *Lactobacillus* isolates and other lactic genera. A significant observation was that APF also played a role in the adhesion of *Lactobacillus* to eukaryotic cells.

For another non-transmissible bacteriophage resistance plasmid, pCI528, a region of DNA that controls its mobilization by conjugation was characterised as was the origin of transfer, *oriT*.

5. New genes

In the course of the programme a series of new genes from the lactic acid bacteria were cloned and sequenced. These include the *dnaK, trp, his, leu-ilv* and *mal* operons of *L. lactis*, β-galactosidase and nuclease genes of *Lactobacillus plantarum*, the S layer gene and the galactose metabolic genes *galK* and *galT* of *Lactobacillus helveticus*.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

- New understanding of the control of gene expression and the development of genetic switches and environmentally controlled promoters.
- Development of food-grade systems for the stable cloning and controlled expression of heterologous genes.
- Characterization of plasmid replication systems and construction of superior cloning and expression vectors.
- Development of sophisticated but practical strategies for integration into the bacterial chromosome.
- Molecular characterization of conjugation processes.

MAJOR COOPERATIVE LINKS

Specific scientific collaborations included work on the characterization and exploitation of lactococcal lactose genes (AFRC;NIZO), plasmid replication (NIZO;UCC) and conjugation processes (PIACENZA;AFRC). In addition, extensive exchanges of data and materials involving all groups took place which are too numerous to specify individually.

Exchanges of personnel included S. Foley (UCC to RUG), J. Law (UCC to RUG), E. Vaughan (UCC to NIZO), V. Bascaran (OVIEDO to AFRC), J.-J. Godon (INRA to AFRC), J. Rodriguez (MADRID to AFRC) and A. Ramos (CTQB to NIZO).

B. PROTEOLYSIS LEAF

OBJECTIVES AND PRIMARY APPROACHES

1. The biochemical and genetic properties of the proteinases and peptidases of lactic acid bacteria
2. The biochemical, bioenergetic and genetic properties of the amino acid and peptide transport systems of lactic acid bacteria
3. The role of individual proteinase and peptidase enzymes or combinations thereof in food manufacturing.

RESULTS AND DISCUSSION

1. Biochemical and genetic production

Proteinases: the maturation of the proteinase from *Lactococcus lactis* was found to proceed via a pre-pro-enzyme. The leader sequence was shown to be removed by a peptidase after which a second protein Prt M associated with the enzyme. Prt M was shown to function as an extracellular molecular chaperone and is essential for the proper folding of the enzyme. Protein engineering of the proteinases has supplied detailed information about the mode of action of these enzymes and has indicated how the activity and/or specificity of the proteinase of *Lactococcus lactis* SK11 can be manipulated. Removal of a segment of the *Lactococcus lactis* SK11 proteinase which is not present in the serine proteinase subtilisin resulted in an enzyme which was still active but had an altered caseinolytic specificity. The transcription of this proteinase was seen to be controlled at the initiation level in which a chromosomally located gene was involved. The sequences in the promoter involved in the medium-induced transcriptional control of the *prtP* and *prtM* genes encoding proteinase production have been identified.
A new serine proteinase was discovered which was shown to be membrane bound. This proteinase (NisP) is coded by one of the genes of the *nis* operon and is the enzyme which splits a leader sequence from the pre-pro-nisin.

The degradation products of β-casein by proteinase of *L. lactis* Wg2 were analysed by HPLC-ion spray mass spectroscopy. About 20 of the peptides formed were small enough to be taken up by the oligopeptide-uptake system. The proteinase of *Lactobacillus bulgaricus* were also studied and purified. It is coded by a gene which is 40% identical to the *Lactococcus lactis* Wg2 proteinase gene. An extracellular phosphatase appeared to play a role in the regulation of the proteolysis of casein.

Peptidases: In the course of the BRIDGE program considerable progress has been made in the identification, purification, characterisation, cloning and sequencing of the peptidases of *Lactococcus lactis*. All enzymes were found to lack a leader sequence which indicates that they are located intracellularly. This is consistent with the results of localisation studies with gold-labelled antibodies against these enzymes. The following peptidases have been characterised: General PepC, aminopeptidases (PepN, Glutamyl, GAP, PepA), tripeptidase (PepT), dipeptidases (PepD, PepR), endopeptidases (PepO, PepF) and carboxyl peptidase (PCP).

Rules for the nomenclature of the peptidases of lactic acid bacteria were formulated. In the BRIDGE program also an extensive analysis of the peptidases from lactobacilli has been carried out. The analysis was mainly at the level of purification and biochemical characterization but the genes encoding several enzymes have also been cloned and sequenced. The enzymes were found to have similar properties to the corresponding enzymes of lactococci but the temperature optima of the *Lactobacillus* enzymes were usually higher.

The genes of the general aminopeptidase (PepN), the X-prolyl-dipeptidyl aminopeptidase (PepX), an iminopeptidase (PepL), a carnosinase (PepV) and two cysteine aminopeptidases (PepC and G) and a prolidase (PepQ) of *Lactobacillus delbrueckii* spp. *lactis* were isolated from *E. coli* transformants.

2. Transport systems

The genes encoding the di/tripeptide (*dtpT*) and oligopeptide (*opp*) transport systems of *Lactococcus lactis* have been cloned and functionally expressed using plasmid vectors. The energetics, substrate specificity and peptide size restriction of the enzymes have been determined. The Opp transport system is thus able to transport several of the peptides formed by the action of the extracellular proteinase on β-casein and supply *L. lactis* with essentially all amino acids needed for optimal growth. Deletion of the gene encoding the substrate binding protein of Opp resulted in a growth-defect in milk. A gene encoding a branched chain amino acid transport system from *Lactobacillus delbrueckii* spp. *lactis* WS 87 has been cloned and sequenced and the functional properties characterised.

3. The role of the individual proteolytic enzymes in casein-hydrolysis and food-manufacturing

When the genes of the endopeptidase (*pepO*) and the tripeptidase (*pepT*) were disrupted, either individually or in combination with *PepN* and *PepX*, no serious growth-defect of *L. lactis* in milk was detected. Since all these enzymes are located intracellularly and no peptides are accumulated intracellularly, these observations indicated that other enzymes with overlapping activities are present. The neutral and alkaline proteinases of *B. subtilis* were expressed in *L. lactis* and the effects on Cheddar cheese ripening have been tested. The presence of the neutral proteinase resulted in a higher water-soluble nitrogen content of the Cheddar cheese and a dramatically accelerated ripening. Cheeses made with strains overexpressing the native proteinase in addition to either *B. subtilis* neutral proteinase or PepN did not ripen in a significantly different manner to controls with the native proteinase. A particular ratio of a *B. subtilis* neutral proteinase-expressing strain to a wild-type commercial cheesemaking starter resulted in a fully-ripened Cheddar cheese in approximately 2 months whose flavour and body/texture qualities exceeded those of the control cheese. A three-fold increase of the lactococcal proteinase had no enhanced effect on casein proteolysis.

Lactococcal strains were also constructed with altered expression of lysyl-aminopeptidases (PepN) by *in vitro/in vivo* recombination of deleted *PepN* and *PepC* genes. Removal of PepN reduced the general aminopeptidase activity by 50% and made the culture low in flavour intensity in cheese trials.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

— The biochemical, functional and genetic properties of the proteolytic enzymes of *Lactococcus lactis* have been determined.

— The biochemical, functional and genetic properties of the peptide transport systems of *L. lactis* have been determined.

— A detailed analysis of β-casein breakdown by proteinase has been performed.

— The localisation of the peptidases, the breakdown products of β-casein by proteinase and the specificities of the peptide transport systems indicated that the only extracellular enzyme of the proteolytic system essential for growth in milk was the proteinase.

— Significant progress has been made in unravelling the proteolytic system of lactobacilli.

— The effects of several strains with a genetically manipulated proteolytic system on cheese ripening have been assessed.

MAJOR COOPERATIVE LINKS

Solid collaborations between the following participants have been established:

IFR, Reading, UK and Chr. Hansen's, Laboratorium, DK on dipeptidase and general aminopeptidase (PepA) and with Univ. Groningen, NL on glutamyl amino peptidase GAPII; the two Groningen groups on the genetic and functional analysis of the peptidases and oligopeptide transport system.
BMF, D and Groningen, NL on growth and proteolytic activities of peptidase deficient mutants of *Lactobacillus delbrueckii* spp. *lactis*.
Groningen, NL and Univ Kaiserslautern, D on the amino acid and peptide transport systems in *Lactococcus* and *Lactobacillus*.
INRA, F and Groningen, NL on the hydrolysis of β casein by proteinase and on proteolytic enzymes of *Lactococcus lactis*.
Extensive exchange of strains, genes, antibodies and other materials has occurred between the participating groups. Joint meetings were held every six months in which the new developments were exchanged in oral presentations, posters and discussion sessions.

C. BACTERIOPHAGE AND PHAGE RESISTANCE LEAF

OBJECTIVES AND PRIMARY APPROACHES

1. Characterization and molecular analysis of phages of lactic acid bacteria(*Lactococcus* and *Lactobacillus* species and *Streptococcus thermophilus*).

2. Study of phage/host interactions in lactic acid bacteria, particularly phage resistance mechanisms and cell surface bacteriophage receptors.

RESULTS AND DISCUSSION

1. Molecular characterization of bacteriophages

The complete nucleotide sequence of the genome of three different lactococcal phages has been determined: the temperate phages R1-T and Tuc2009 and the virulent prolate-headed phage bIL67. The chromosome of phage R1-T is a 33.3 kb linear double stranded DNA molecule with cohesive ends. Of the 49 open reading frames that were present, only three were orientated in a direction opposite to that of the others. The phage attachment site and the gene coding for the integrase were identified. Furthermore, it has been shown that integration of phage R1-T occurred at a unique site in the chromosome of its host, *Lactococcus lactis* subsp. *cremoris*.

The phage Tuc2009, isolated from *Lactococcus lactis* subsp. *cremoris* UC509, possesses a double stranded DNA genome with an estimated length of 38.7 kb, which is packaged into the phage head by a so-called headful mechanism. Sequence analysis of the genome revealed a putative repressor-encoding gene *cI*, presumably involved in the establishment of lysogeny, as well as two genetic determinants (encoding a lysin and a putative holin) required for host cell lysis. Furthermore, the gene encoding the integrase was identified in addition to the attachment site of Tuc2009, which enabled the construction of a site-specific integration system. The determination of the N-terminal sequence of several Tuc2009 phage proteins permitted the identification and localisation of two structural genes on the phage genome.

The genome (22.1 kb) of the prolate-headed phage bIL67 was found to contain 37 ORFs organized in two divergent transcriptional units. Putative functions have been assigned to six of the ORFs: a DNA polymerase, a protein involved in recombination, a lysin, a terminase subunit, a structural protein and a holin.

Several lysins of lactococcal virulent phages were also characterized genetically and biochemically; those of prolate-headed phages P001 and ØML3 are muramidases and are highly homologous whereas the lysin of the small-isometric phage phiUS3 appeared to be an amidase.

TP-J34 is a temperate isometric-headed phage which is inducible from the host strain *Streptococcus thermophilus* J34. 99% of the phage particles released from the lysogenic host strain were found to be defective (tail-less). Phage TP-J34 was shown to be closely related to one subclass of virulent phages. The phage genome (47 kb) was demonstrated to be linear, terminally redundant and circularly permuted. The *pac*-site and the *attP*-site were localized. Based on the physical map, 85% of the phage DNA could be cloned in *E. coli* vectors. When TP-J34 was propagated lytically on a prophage-cured indicator strain, a site-specific deletion occurred in the phage genome. In this non-coding region, sequence analysis revealed several direct AT-rich repeats.

A temperate *Lactobacillus casei* phage, Ø393-A2, was shown to possess a 47 kb genome with cohesive ends. A physical map of its DNA was constructed and its general characteristics determined. Furthermore, several functions have been localized in its genome including the integration region, the repressor and the genes coding for the major capsid protein and the terminase, which was found adjacent to the cohesive ends. Detailed deletion analysis has shown that a region of at least 4 kb of the phage genome was dispensable for lytic development.

Significant progress has also been made in the molecular characterization of the *Lactobacillus delbrueckii* temperate phage mv4. Three regions involved in specific

functions and having a biotechnological interest have been characterized by sequencing more than one third of the phage genome. On the gene cluster encoding the cell lysis activity (2.8 kb), the *lysA* gene coded for a protein (LysA), which could be a muramidase since its N-terminal part showed good homology with *Chalaropsis*-type muramidase. The size and the activity of LysA was confirmed by enzymography on SDS-PAGE. The *lysA* gene was part of an operon and was cotranscribed with the *lysB* gene. LysB could represent the holin protein involved in LysA translocation across the cell membrane. Upstream of the *lysBA* genes, a putative late promoter sequence has been identified. In a similar manner, the genetic organization of the region (3.3 kb) encoding the main capsid protein (34 kDa) of mv4 has been characterized; the *g34* gene was located between a gene coding for a 20 kDa immunoreactive protein and three ORFs of unknown function. No phage promoter could be detected upstream of the *g34* gene by primer extension analysis, suggesting that this gene was part of an operon.

Site-specific integration of the mv4 genome into the $tRNA^{adr}$ gene of the host chromosome was demonstrated. The cloning and sequencing of the phage *attP* surrounding region (3.2 kb) allowed the genetic characterization of this early region, particularly the ORF coding for the mv4 integrase. The host-phage junction fragments *attL* and *attR* as well as the bacterial attachment site *attB* have been sequenced. The 'core' sequence where the strand exchange takes place is 17 bp long and overlaps the 3' end of a $tRNA^{adr}$ gene.

To better understand the relationships between virulent and temperate phages, these data have been compared to those obtained with the related virulent phage LL-H. The genes involved in late functions (*lysA* and *g34*) appeared to be subjected to a strong selective pressure since they were found to be highly conserved between the mv4 and LL-H phages whereas early regions were completely divergent, except for the presence of a remnant *attP* site and of a truncated int gene on phage LL-H genome. The temperate origin of this virulent phage was demonstrated. The molecular taxonomy of *Leuconostoc oenos* and *mesenteroides* phages revealed a great genetic diversity among these agents.

2. Study of phage/host interactions in lactic acid bacteria.

2.1 Phage resistance mechanisms

A 46 kb absorption-blocking plasmid, pC1528, was found to alter the cell surface characteristics of its host, *L. lactis* subsp. *cremoris* UC503, by directing the production of a galactose and rhamnose containing hydrophilic polymer.

Three distinct lactococcal R/M systems have been characterized. Plasmids coding for the two different type II R/M systems, *LlaAI* and *LlaBI*, have been isolated from the *Lactococcus lactis* subsp. *cremoris* strains W9 and W56. The recognition sequences of the restriction endonucleases were determined and the genes responsible for these R/M systems were cloned and sequenced. The *LlaAI* R/M system, composed of one endonuclease- and two adenine methylase- encoding genes, exhibited significant homology to the *DpnII* system of *S. pneumoniae*. *LlaBI* endonuclease was shown to be an isoschizomer of *SfcI*.

The genes of the chromosomally located *ScrFI* R/M system from *L. lactis* subsp. *cremoris* UC503 have been analysed. Two genes, which flank the *ScrFI* endonuclease gene, encode 5-methylcytosine methylases which independently confer resistance to *ScrFI* digestion. The variable region of these methylases, which contains the target recognition domain and which is also responsible for determining which cytosine is methylated, were found to be quite different. This indicated

that they may have alternative substrate or methylation specificities. These genes have been reintroduced and expressed in *Lactococcus*. Furthermore, genes encoding 50s ribosomal proteins were observed upstream of the *ScrFI* R/M determinants.

An R/M mechanism was found in spontaneous phage resistant mutants of *Streptococcus thermophilus* To39. Expression of this phage resistance mechanism was found to be the result of a cointegration event involving two resident plasmids. The resulting cointegrated plasmid of 9.6 kb was conjugally co-mobilized into a number of *S. thermophilus* strains. These strains were successfully used for some pilot plant productions.

Five Abi determinants (*abi-416, abi-420, abi-105, abi-750* and *pNP40* have been cloned and were shown to be different on the basis of phenotypic and genetic criteria.

Three distinct phage resistance mechanisms encoded by plasmid pNP40 have been identified. Two of these, AbiD and AbiE, mediate resistance via abortive infection. Two ORFs appear to be required for AbiD, whilst a single ORF is sufficient for the expression of AbiE. An ORF encoding a RecA-like protein was shown to be located in the intervening region between *abiD* and *abiE*, but has not been implicated in the phage resistance phenotype. In addition, a novel phage resistance mechanism on pNP40 operates by preventing phage DNA injection.

The Abi750 abortive infection mechanism, encoded by the conjugative plasmid pCI750, manifests itself by conferring complete resistance to small isometric-headed phages and partial resistance to prolate-headed phages. This mechanism acts late in the phage lytic cycle. Nucleotide sequence analysis of pC1750 has revealed that more than one ORF may be responsible for the Abi750 phenotype.

Two further Abi determinants *abi*-416 and *abi*-105 have been sequenced and the proteins they encode were found not to share homology with proteins from the data banks. The target sites of *abi*-420, *abi*-105, and *abi*-416 have been localized in different parts of the phage genome and are expressed at the early, middle and late period of the phage infection cycle, respectively. Two of these targets have been cloned and sequenced. The *abi*-416 target on phage blL41 corresponds to an operon coding for structural proteins whereas the *abi*-105 target on phage blL66 corresponds to an operon of unknown function. It was shown that phage mRNAs were degraded in the presence of *abi*-416, suggesting that the abi gene product activated a phage encoded RNAase.

The Abi system encoded by plasmid p1149-3 of *L. lactis* subsp. *cremoris* 1149 was shown to mediate resistance to a variety of mostly isometric-headed phages by blocking a step after phage DNA replication.

2.2 Cell surface receptors

The surface located substance acting as a receptor for the virulent phage CNRZ 832 was identified in the surface layer of the bacterial host *Lactobacillus helveticus* CNRZ 892. This protein was isolated, purified and reconstituted *in vitro* and used to inactivate phage particles. It is the first time that it is possible to assign a functional role to a *Lactobacillus* surface layer.

MAJOR SCIENTIFIC BREAKTHROUGHS

— Nucleotide sequence determination of the entire genome of the lactococcal temperate phage Tuc2009 and the virulent prolate-headed phage blL67.

— Characterization of several bacteriophage functions, i.e. cell lysis, site-specific recombination, morphogenesis, allowing the development of purpose designed food grade vectors.

— Demonstration of the temperate origin of a virulent phage.

— Molecular analysis of three lactococcal type II R/M systems: *LlaA1, LlaBI* and *ScrFI.*

— Localization and characterization of Abi targets on the genome of lactococcal phages.

— Combination of R/M systems with Abi mechanisms resulting in strains with a higher degree of phage resistance.

MAJOR COOPERATIVE LINKS

Exchange of material (strains, phages and genes) and joint experiments were performed between the groups working on the molecular characterization of the phages (CNRS, BMF, AMO, INRA, UCC) as well as those involved in the characterization of the phage resistance mechanisms (INRA, UCC, Royal Vet. Agr. Univ. Frederiksberg, BMF, CUP)
Exchanges of researchers: S. Foley (UCC) 14 months at Groningen; J. Law (UCC) post-doctoral position at Groningen, NL. E. Vaughan (UCC) post-doctoral position at NIZO; N. Filser (Kaiserslautern) 9 months at CNRS Toulouse.

D. ANTIMICROBIALS LEAF

BACKGROUND INFORMATION

Bacteriocins are proteinaceous compounds produced by bacteria and are inhibitory against other bacteria. Bacteriocins from the food-grade lactic acid bacteria could potentially be used for natural food preservation and the control of complex microbial ecosystems such as fermented foods/feed or the gastro-intestinal tract. In addition, bacteriocin production or immunity could be used as a food-grade marker in recombinant-DNA technology.

OBJECTIVES AND PRIMARY APPROACHES

1. Isolation of bacteriocin-producing lactic acid bacteria with interesting inhibitory spectra (including pathogens and spoilage bacteria).
2. Purification and characterization of the bacteriocins.
3. Localization, isolation and characterization of the genetic determinants involved in bacteriocin production and immunity.
4. Optimization of bacteriocin production by physiological studies and recombinant DNA technology.
5. Study of the application potential of the bacteriocins.

RESULTS AND DISCUSSION

1. Screening for bacterocins

Over 3000 strains of lactic acid bacteria were isolated from various European sources, including local (fermented) food products such as meat, fish, dairy products, vegetables and wine and screened for bacteriocin production. Approximately 80 positive strains were found and ca. 30 bacteriocins were studied in more detail. In order to exclude false positive results and to determine if identical bacteriocins had been studied by several participants, a standardized method employing a standardized panel of target strains (both lactic acid bacteria as spoilage organisms and pathogens) was used. Comparison of the antimicrobial spectrum of the various bacteriocins indicated that most of the selected compounds had different inhibitory activities. Bacteriocin production was found within different genera: lactococci, pediococci, carnobacteria and lactobacilli. The inhibition spectra varied from extremely narrow (lactococcin A) to rather broad (lacticin 481, plantaricin S), while bacteriocins with high activity against *Listeria* (pediocin PA- 1) or *Clostridium* (acidocin B, lactacin 247) were also identified. Some strains appeared to produce more than one bacteriocin at the same time.

Later additional screenings, in some of which PCR analysis and genetic probes were used, revealed that lactocin S producing lactobacilli and pediocin PA-1 producing pediococci are quite widespread, since they could be isolated in different countries from various sources. Also, nisin producing lactococci could be isolated from meat environments.

2. *(a) Purification/characterization of bacteriocins*

Several of the bacteriocins studied within the BRIDGE programme have been purified. Generally, the purification procedure employed ammonium sulphate precipitation or butanol extraction as a first step, followed by ion exchange chromatography, hydrophobic interaction chromatography and reverse phase-HPLC. Biochemical characterization revealed that novel representatives of different bacteriocin 'classes' were studied:

(I) Lantibiotics, which are small peptides containing unusual amino acids such as lanthionine, e.g. lactacin 481, lactocin S;
(II) Small, non-lanthionine containing peptides, e.g. lactococcins A/B/M/G, pediocins PA-1/L50, plantaricins C/F/S/T, acidocin B and lactacin 247.

Some of these bacteriocins were most active as a two-peptide system (lactococcins M/G, plantaricin S), whereas lactococcin B is only active in a reduced form.

2. *(b) Mode of action of bacteriocins*

Generally the bacteriocins from lactic acid bacteria exert a bactericidal and sometimes lytic effect on sensitive organisms. It was concluded that the degree of sensitivity appeared to be affected by the physiological state of the target strain (log phase vs. stationary phase), the amount of bacteriocin used and the medium composition. More mechanistic mode of action studies with lactococcins, pediocin PA-1 and plantaricins C/F showed that these bacteriocins inhibited sensitive cells by forming pores in the cytoplasmic membrane. As a result, the transmembrane pH gradient and membrane potential are dissipated and efflux of essential cellular metabolites occurs. The size of the pores increases with increasing bacteriocin concentration.

3. *(a) Genetics of bacteriocin production*

In many cases the genes responsible for bacteriocin production (and immunity) are plasmid-located. Several genes have been cloned and (over-)expressed in other lactic acid bacteria. For different bacteriocin producers, plasmid cured Bac^- derivatives have become available. The genetic organization of lactococcin A production has been studied in most detail. The genes essential for bacteriocin activity were found to be organized in an operon and were involved in precursor synthesis, processing, excretion and immunity.
DNA sequencing of the structural genes for 'class II' bacteriocins revealed that most shared a common and characteristic, i.e. a $GlyGly^{-1+1}X$, processing site. However, acidocin B appeared to be an exception, since a more general processing site obeying the von Heyne rule is used.

3. *(b) Host Immunity*

Production of an immunity factor is essential for host-resistance. The immunity protein for lactococcin A has been purified. Treatment of sensitive cells with immunity protein could not prevent killing by the bacteriocin, indicating that resistance did not result from a direct binding of immunity protein to lactococcin A. Additional experiments using monoclonal antibodies showed that the immunity protein was predominantly membrane bound and probably prevented the action of the bacteriocin by shielding a receptor molecule or preventing penetration of the receptor/bacteriocin complex into the cytoplasmic membrane. Upon introduction of the genes responsible for acidocin B production and immunity into *Lb. fermentum*, this organism also acquired resistance against plantaricin F although it remained sensitive to nisin and some other bacteriocins. Similar results were obtained with an acidocin B resistant mutant of *Lb. fermentum*. Future studies on bacteriocin structure, (cross)immunity, resistance/sensitivity and inhibitory spectrum will be essential for the elucidation of the structure/function relationship.

4. Bacteriocins production

Plantaricins S and T are produced by one strain of *Lb. plantarum*, during the log and stationary phase, respectively. Production was found to be optimal in the presence of 4% NaCl in a fermentor without pH control. The conditions under which bacteriocin production occurred were investigated in more detail for pediocin PA-1, bavaricin and acidocin B. The optimum conditions of pH and temperature were established in fermentation studies. Pediocin PA-1 and acidocin B were produced during the entire exponential phase while bavaricin was produced only in the early log phase and decomposed thereafter. Production of acidocin B could also be achieved by washed cells in a simple, well defined medium which did not support growth. This facilitated subsequent purification.

5. Application studies

The effect of plantaricin S on other lactic acid bacteria and spoilage organisms was studied both in model experiments and in natural olive fermentations. When inoculated into olive brine, the plantaricin S producer (genetically marked and selectively enumerated) proliferated to dominate the epiphytic flora and persisted throughout the fermentation. In contrast, a Bac^- derivative strain could not be isolated a few days after inoculation. Model experiments also revealed that bacteriocin production was one of the mechanisms which allowed this strain to co-exist with more salt tolerant lactobacilli in olive brine.

After studying the interaction of acidocin B with starter cultures, milk and clostridia *in vitro*, this bacteriocin was used in model cheesemaking experiments. Clostridial activity could be delayed but not inhibited completely. A shift from acidocin B-sensitive to acidocin B-resistant clostridia was observed during cheese ripening, suggesting that either adaptation or accumulation of insensitive strains had occurred.
Application of the pediocin PA-1 producing *Pediococcus* in simulated meat systems and in sausage production resulted in a significant restriction of *Listeria monocytogenes* (1000-fold reduction in numbers during the experiment). This effect was not caused by formation of lactic acid, since no inhibitory effect was observed when a Bac$^-$ variant of the *Pediococcus* was used. In addition, production of pediocin PA-1 in these food products could be demonstrated. A notable result was that oral addition of lactacin 247 producing *Lb. acidophilus* resulted in a decrease in the number of *Clostridia* in the gastrointestinal tract of mice, in comparison to addition of the *bac*$^-$ mutant.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

- A large number of novel bacteriocins with interesting antimicrobial spectra have been identified.
- The role of bacteriocin-producing lactic acid bacteria in controlling food fermentations was established. Strains producing the same bacteriocins could be isolated from various locations in Europe from fermented foods, suggesting that these strains are dominant in natural ecosystems.
- Fundamental knowledge of the biochemical structure of bacteriocins, the genes involved in production, and the mode of action was obtained.
- A beginning was made in elucidating the relationship between structure and function. Ultimately this could result in the application of tailor-made bacteriocins with enhanced properties, combinations of bacteriocins and/or improved starter cultures for the production of safer foods.

MAJOR COOPERATIVE LINKS

The BRIDGE programme has allowed extensive collaboration among the participants within the 'antimicrobial LEAF'. Exchange of knowledge, materials and personnel occurred frequently. In particular, participants who are experts on certain topics (purification, mode of action, genetics) were consulted on numerous occasions; this resulted in a number of joint projects and joint, multinational publications.

E. METABOLISM LEAF

BACKGROUND INFORMATION

Metabolic conversions that are initiated by lactic acid bacteria are of primary importance in the industrial fermentation of milk, meat, sourdough or wine. Their main role is to contribute to the preservation, flavour and texture of the fermented foods. Whilst in many cases the substrates that are transported and metabolized are known, our knowledge about the metabolic pathways, their enzymology, or energetics is limited or non-existent. In addition, the genes involved in the key metabolic processes and their control have not been characterized. This seriously

limits the potential that modern fermentation technology and metabolic engineering offers in improving industrial food fermentations.

OBJECTIVES AND PRIMARY APPROACHES

1. Transport, energetics and pathway analysis of important fermentations.
Various selected fermentative conversions have been studied in lactic acid bacteria involved in the production of wine, sourdough, meat, and dairy commodities. This has been realized by integrating physiological, enzymological and spectroscopical (*in vivo* NMR) approaches. The main focus of attention has been the metabolism of sugars (glucose, maltose, lactose and galactose), citrate and malate.

2. Genetic characterization, expression and regulation of metabolic pathways.
Major attention has been given to the lactose and galactose operons in dairy lactobacilli and lactococci and the genes involved in the lactococcal pathway leading to the diacetyl precursor α-acetolactate. A combined genetic, physiological and biochemical approach has been taken and in selected cases protein engineering has been applied.

3. Engineering of metabolic pathways.
Major attention has been focused on engineering pyruvate metabolism of lactococci in order to increase the metabolic potential of these important dairy starters. An integrated genetic and physiological approach has been taken.

RESULTS AND DISCUSSION

1. Transport, energetics and pathway analysis of important fermentations

Sugar, citrate and malate metabolism was studied in various strains of *Leuconostoc oenos* that are involved in the malolactic fermentation in wine. The mechanism and bioenergetics of transport of citric and malic acid were elucidated. In addition, the inhibitory effect of glucose on malolactic fermentation was investigated. By using *in vivo* NMR, a novel pathway for erythritol production was discovered and its regulation was established. Finally, the co-metabolism of sugar and citric acid was investigated and the pathway for diacetyl production was elucidated in *Leuconostoc oenos*.

In the sourdough starters *Lactobacillus sanfrancisco* and the novel isolate *Lactobacillus pontis* (named after the BRIDGE project; see report of the Screening LEAF) the metabolism of glucose and maltose was studied. The mechanism and energetics of maltose uptake and glucose excretion was elucidated in *Lactobacillus sanfrancisco*. The co-fermentation of substrates was studied in this heterofermentative *Lactobacillus* with specific attention directed to energy production and co-factor regeneration.

The survival, stress response and product formation were studied in the principal lactobacilli used in meat fermentations, *Lactobacillus* species *sake* and *curvatus*. The production of bacteriocin by *Lactobacillus curvatus* is an important factor affecting its competitive capacity and survival in sausage fermentations.

The metabolism of citrate in the dairy starter cultures *Leuconostoc* and *Lactococcus* spp. was investigated. The pathway for diacetyl production in *Lactococcus lactis* subsp. *diacetylactis* has been established by enzymatic and *in vivo* NMR studies and involves a non-enzymatic conversion of α-acetolactate into diacetyl. Citrate transport, degradation, and co-metabolism was also studied in *Leuconostoc lactis* spp. The key enzyme, acetolactate synthase, was partially purified from *Leuconostoc lactis* and purified to homogeneity from *Lactococcus lactis*. In addition, the impor-

tant enzyme, acetolactate decarboxylase, was purified from *Lactococcus lactis* and its allosteric control by amino acids was established. Finally, sugar metabolism in *Lactobacillus helveticus* was studied and glucose repression of lactose during galactose utilization was observed.

2. Genetic characterization, expression and regulation of metabolic pathways

The *lac* and *gal* operons appear to be linked in the chromosome of *Lactobacillus helveticus*. The instability of the capacity to utilize lactose was studied in this thermophilic *Lactobacillus* and suggested two different mechanisms.
Lactose metabolism and its control was also studied in detail in *Lactococcus lactis*. The *lacR* gene encodes the repressor of the *lacABCDFEGX* operon which encodes lactose transport and galactose degradation. Two LacR operator sites have been identified by DNAse I footprinting. Protein engineering studies of LacR using an alanine scanning approach have located the recognition helix and provided mutants with improved operator binding. Tagatose-6-phosphate has been identified as the molecular inducer of *lac* operon expression. Protein engineering studies have identified the charged residues in LacR involved in response to the inducer. Catabolite repression was found to constitute an additional control system and a glucose responsive element was identified in the *lac* promoter region. A model for the regulation of *lac* operon gene expression has been developed and the first generation of inducible expression vectors have been developed for lactic acid bacteria.

3. Engineering of metabolic pathways

The metabolic engineering of pyruvate metabolism in lactococci was initiated by the cloning and characterization of the *L. lactis ldh* gene followed by the construction of LDH-deficient mutants by replacement recombination. Remarkably, the constructed LDH-deficient strains were viable, and even displayed an increased cell yield in batch culture, while showing a mixed acid fermentation.

As a first example of metabolic engineering the *Zymomonas mobilis* pyruvate decarboxylase and alcohol dehydrogenase genes were expressed as gene fusions in *L. lactis* and resulted in alcohol overproduction from lactose especially in LDH-deficient strains.

A second industrially relevant metabolic engineering activity involved the rerouting of pyruvate to the diacetyl precursor α-acetolactate. For this purpose the *als* and *ilyBN* genes were cloned and overexpressed in several *L. lactis* strains. Both genes encode enzymes involved in the production of α-acetolactate but differ in their affinities for pyruvate. In LDH-deficient strains the overexpression of the *als* gene resulted in a major shift of the flux from lactose towards the flavour precursor α-acetolactate and acetoin/butanediol, that constituted quantitatively the most important end products.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

- Molecular characterization of *L. lactis lac* operon control and identification of the physiological inducer.
- Increased production in *Lactococcus* of ethanol and acetoin by metabolic engineering.

— Construction of LDH-deficient lactococci that show higher yields and do not produce lactic acid.

— Determination of important metabolic pathways by using *in vivo* NMR.

— Elucidation of the energetics of the transport of sugars, citrate and end products.

— Improved prediction of performance and handling of meat and sourdough starter cultures.

MAJOR COOPERATIVE LINKS

The groups participating in the Metabolism LEAF held meetings at six monthly intervals where there was exchange of data, ideas and know-how. In addition, there was exchange of personnel between laboratories and joint publications.

F. SCREENING LEAF

BACKGROUND INFORMATION

The development of new starter cultures with improved qualities [enhanced peptidolytic and esterolytic activity, production of exopolysaccharides (EPS)] is based on the screening of large numbers of new isolates. The most promising strains should be selected and identified in order to confirm their status as food-grade microorganisms.

Correct identification of lactic acid bacteria is a prerequisite to their use as food-grade organisms in food fermentations. Traditional phenotypic methodologies, although necessary for descriptions, have often been turned into less reliable classification schemes. Modern molecular taxonomic methods (rRNA sequencing, protein fingerprinting using SDS-PAGE) should therefore be applied on the wide variety of lactic acid bacteria in order to verify and improve the existing classification of these organisms. The application of advanced computer algorithms on the large data sets obtained by these methods should improve the existing classification and identification tools.

OBJECTIVES AND PRIMARY APPROACHES

1. Isolation of new microorganisms from a large pool of natural food/feed fermentation products (such as fermented sausages, fermented olives, sauerkraut, sourdough, dairy products) in order to improve their quality (safety, shelf life, organoleptic quality, flavour).

2. Conservation of important strains.

3. Screening for important phenotypic properties (EPS production and enhanced activities of amino- and diaminopeptidases and esterases) in LAB.

4. Identification of the most promising new isolates using traditional phenotypic-tests verified by the use of molecular taxonomic methods (rRNA sequencing,DNA:DNA hybridisation and protein fingerprinting using SDS-PAGE) in order to confirm their food-grade status.

5. Improvement of classification schemes by including large numbers of strains using a thorough computer-based comparison of the newly obtained data with the available information of well-known reference strains.

6. Design of reliable and fast identification and detection tools (specific rRNA-targeted oligonucleotide probes), to be applied *in vitro* (using different hybridisation methods) or *in situ* (using fluorescent-labelled oligonucleotideprobes).

7. Further molecular characterisation of the EPS and localization of their encoding genes.

RESULTS AND DISCUSSION

1. Isolation of strains, preparation of a collection of LAB cultures

More than 600 strains have been isolated from traditional European food products and processes (Unilever Research Laboratories). A collection of 150 strains of traditional Greek cheese and yoghurt and 181 strains from Kasseri cheese were made available for screening and/or taxonomic analysis (Agricultural University Athens). A total of 952 strains of various LAB genera were received at the Laboratorium voor Microbiologie Gent (80% were offered by 11 T-project participants; 20% were deposited by other collaborating laboratories). A collection of about 1400 reference strains of LAB (LMG Culture Collection, Laboratorium voor Microbiologie Gent) was made available to the project. This large collection of bacteria was essential for the planned screening activities and for taxonomic analysis.

2. Culture collection maintenance.

All strains received at the Laboratorium voor Microbiologie Gent were verified for purity by plating and microscopic examination and were deposited in the closed LAB culture collection. About 97% of the strains were lyophilised and/or preserved in liquid nitrogen; remaining strains were destroyed on request or lost prior to analysis. All requests for strain exchange have been granted by the respective depositors. After publication of the results, 84 strains have been transferred to the public LMG culture collection of the Laboratorium voor Microbiologie Gent.

3. Screening for important phenotypic properties.

Thirty EPS producing strains were selected out of more than 600 new isolates. The rheological properties were investigated, and in collaboration with a specialised laboratory, the structure of some of the EPS's has been determined. A patent on the food application of one very promising EPS was filed. Amino- and diaminopeptidases activities (eight substrates) were registered for almost 500 strains of LAB; esterase activities (three substrates) were measured in 150 strains from traditional Greek dairy products. For the occurrence of both enzymes, a considerable heterogeneity was observed between the different genera and sometimes between individual species. Between taxonomically related strains, the activities of the specific enzymes varied considerably.

4. Identification of new isolates.

Using morphological, physiological and nutritional tests, 65 strains were identified to species level and 266 strains were partially characterized at the Agricultural University of Athens. For 150 strains these phenotypic results were already verified by SDS-PAGE of whole-cell proteins; the other stains are at present being analysed. More than 93% of the 952 strains received at the Laboratorium voor Microbiologie Gent, were identified to the species level by SDS-PAGE of whole-cell proteins, although often organisms were received which were difficult to identify by traditional methods. The remaining strains were (i) shown to be repre-

sentatives of new species, (ii) shown not to belong to the LAB, or (iii) are still under investigation. Enhanced identification of LAB by SDS-PAGE could only be accomplished by

(i) analysing large numbers of new field strains,
(ii) the extension and optimisation of the database of protein patterns of reference strains, and
(iii) the use of automated data-comparison in order to obtain the maximum possible information from this large data set (creation of a library of normalised protein pattern fingerprints and implementation of new algorithms for data comparison).

SDS-PAGE identification results were often verified by the use of other molecular taxonomic methods (rRNA sequencing, DNA:DNA hybridisation, oligonucleotide probe hybridisations) performed at the Technical University Munchen (see below).

5. Improvement of classification schemes

Classification schemes could be improved by comparing the phylogenetic and taxonomic information obtained on large numbers of newly analysed strains with existing classification schemes. New groups of organisms, detected by SDS-PAGE or by DNA:DNA hybridisations were phylogenetically analysed by 16S rRNA sequencing (Technical University Munchen) and were further phenotypically analysed in order to describe them properly. As a result the two subspecies of *Streptococcus salivarius (salivarius and thermophilus)* were elevated to species level, a new *Lactobacillus* species (*Lactobacillus pontis*, named after the BRIDGE project) was described, the description of a new subspecies of *Lactobacillus hilgardii* is being prepared, and the possibility of phenotypically describing a new species of the genus *Enterococcus* is currently being examined. Interesting results were also obtained for *Lactobacillus acidophilus* (which contained at least six different protein electrophoretic groups), *Lactobacillus lactis* subsp. *lactis* and subsp. *cremoris* (120 strains could be differentiated by SDS-PAGE), the *Lactobacillus casei* complex (confirming the findings of Collins *et al.*, 1989), *Lactobacillus brevis* (which is consistently being mis-identified due to high phenotypic similarities with several other *Lactobacillus* species) and *Lactobacillus sake* (which is very often confused with *Lb. curvatus* and *Lb. bavaricus*).

6. Design of fast identification and detection tools

For more than 30 different species of LAB, 16S and/or 23S rRNA-based oligonucleotide probes were developed and evaluated (Technical University Munchen). SDS-PAGE results were often useful to select the strains submitted to comparative sequence analysis and the strains necessary to test the specificity of the newly designed probes. An excellent agreement was found between both techniques. Various species-specific probes were successfully applied on some misclassified strains and applied as captured probes in a non-radioactive reverse dot blot hybridization reaction. Colony hybridization with rRNA-targeted oligonucleotides was improved. Fluorescent-labelled oligonucleotide probes were used for the identification and *in situ* detection of lactococci and certain enterococci by whole-cell hybridization. Moreover, the use of *in vitro* rRNA gene fragment amplification enabled direct detection of LAB, often with species identification, within one day.

7. Molecular characterisation of the EPSes; search for their encoding genes

The rheological properties of seven of the EPSes were determined as was the structure of the repeating unit for five of them. One EPS, produced by a *Lactobacillus sake* strain, might be very well suited for application as a food-waterphase stabilizer. The factors which enhanced production were investigated and optimized and at present a search is being made for the genes involved in the biosynthesis of EPS. Several strategies, including the use of DNA probes of homologous genes in *Lactococcus* (L-LDH gene) and *Salmonella, Rhizobium, Xanthomonas* (acetylase gene) are being used.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

- For several genera and species it has been shown that correct identification by traditional phenotypic tests is impossible or not reliable. Classification schemes can be updated by using modern molecular taxonomic techniques such as SDS-PAGE and oligonucleotide probe hybridisations.
- An excellent agreement was found between identification results obtained with rRNA targeted oligonucleotide probes and by SDS-PAGE of whole-cell proteins.
- Isolation of an EPS producing *Lactobacillus sake* strain which might be very well suited for application as a food-waterphase stabilizer.
- The gene encoding L-LDH of *Lactobacillus sake* was isolated and can be used to study the influence of lactate on EPS production.
- A patent on the food application of *Lactobacillus sake* EPS was filed.

MAJOR COOPERATIVE LINKS

Three contractor meetings and three subLEAF meetings were held.
Isolation of South European LAB strains (Unilever Research Laboratory, Vlaardingen, NL and Centro de Tecnologia Quimica e Biologica, Oeiras, P), exchange and characterization (identification) of more than 1300 strains (Laboratorium voor Microbiologie Gent, B and the Agricultural University Athens, GR for 10 different participants), central 16S rRNA sequencing (Technical University Munchen, D for the Agricultural University Athens, GR; Universitat Hohenheim, D; Escola Superiora de Biotecnologia, Porto, P; and the Bundesanstalt fur Milchforschung, Kiel, D).

Joint screening efforts (exchange and uniformisation of procedures) for more than 650 strains (Unilever Research Laboratory, Vlaardingen, NL; Agricultural University Athens, GR; Laboratorium voor Microbiologie Gent, B) and training activities for SDS-PAGE techniques and numerical analyses (Laboratorium voor Microbiologie Gent for Agricultural University Athens, GR; Escola Superiora de Biotecnologia, Porto, P; Universidade de Tras-Os-Montes e Alto Douro, Vila Real, P; Universitat Hohenheim, D).

PUBLICATIONS

Joint publications

Adebawo, O.O., J.L. Ruiz-Barba and P. J. Warner. (1993). Development of lysine overproducing strains of *Lactobacillus plantarum* for improvement of the nutritional value of fermented foods. FEMS Microbiol. Rev. **12**: 166.

Anderson, S.V., Collins, M.A. and Poolman, B. (1994). Characterisation of sugar transport mechanisms in *Lactobacillus helveticus*. (in preparation).
Arendt, E.K., G.F. Fitzgerald, A. Coffey and W.P. Hammes. (1994). Bakteriophagen den Kampf ansagen. Lebensmitteltechnik **6**: 40-44.
Boizet, B., M. Mata, O. Mignot, P. Ritzenthaler and T. Sozzi. (1992). Taxonomic characterisation of *Leuconostoc mesenteroides* and *Leuconostoc oenos* bacteriophages. FEMS Microbiol. Lett. **90**: 211-216.
Brurberg, M.B., A.J. Haandrikman, K.J. Leenhouts, G. Venema, and I.F. Nes. (1994). Expression of a chitinase gene from *Serratia marcescens* in *Lactococcus lactis* and *Lactobacillus plantarum*. Appl. Microbiol. Biotechn. (in press).
Callegari M.L., Séchaud Laurent, M. Rousseau, V. Bottazzi, L. Morelli and J.P. Accolas. (1994). The S-layer protein of *Lactobacillus helveticus* CNRZ 892 contains the receptor for virulent phage 832-B1. (submitted).
Chikindas, M.L., M.J. Garcia-Garcera, A.J.M. Driessen, A.M. Ledeboer, J. Nissen-Meyer, I.F. Nes, T. Abee, W.N. Konings, and G. Venema. (1993). Pediocin PA-1, a bacteriocin from *Pedococcus acidilactici* PAC1.0, forms hydrophilic pores in the cytoplasmic membrane of target cells. Appl. Environ. Microbiol. **59**: 3577-3584.
Cintas, L.M., J.M. Rodriguez, M.F. Fernandez, H. Holo, K. Sletten, P.E. Hernandez and I.F. Nes. (1994). Isolation and characterisation of pediocin L50, a new bacteriocin produced by *Pedococcus acidilactici* L50. Appl. and Environ. Microbiol. (in preparation).
Cocconcelli P.S., M.J. Gasson, L. Morelli and V. Bottazzi. (1991). Single stranded DNA plasmid, vector construction and cloning of *Bacillus stearothermophilus* α-amylase in *Lactobacillus*. Res. Microbiol. **142**: 643-652.
Couto, J.A., B. Pot, M. Ehrmann, K.H. Schleifer, K. Kersters and T. Hogg. (1994). Taxonomic analysis of *Lactobacillus hilgardii* strains from wine source by means of numerical analysis of protein electrophoretograms and genotypic methods. (in preparation).
Curk, M.C., J.M. Boeufgras, B. Decaris, F. Gavini, K. Kersters, J.P. Larpent, P. Le Bourgeois, P. Renault, H. de Roissart, and C. Rouvier. (1994). Methodes d'identification des bacteries lactiques. In Bacteries Lactiques I.H. de Roissart, and F.M. Luquet (eds) pp. 141-168. Lorica, Uriage, France.
Davis, R., D. van der Lelie, A. Mercenier, C. Daly and G.F. Fitzgerald. (1993). ScrFI Restriction-modification system of *Lactococcus lactis* subsp *cremoris UC503:* Cloning and characterisation of two *Scr*FI methylase genes. Appl. Environ. Microbiol. **59**: 777-785.
Descheemaeker, P., B. Pot, A. Ledeboer, T. Verrips and K. Kersters. (1994). Identification of *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris* by SDS-PAGE analysis. Syst. App. Microbiol. (in press).
De Vos, W.M., J. Huis in 't Veld and B. Poolman (Eds.) (1993). Proceedings of the Fourth Symposium on Lactic Acid Bacteria. FEMS Microbiol. Rev. **12**: 1-272.
Devriese, L., B. Pot and M.D. Collins. (1993). Phenotypic identification of the genus *Enterococcus* and differentiation of phylogenetically distinct enterococcal species and species groups: A review. J. Appl. Bacteriol. **75**: 399-408.
Dupont, L., A. Vasala, M. Mikkonen, B. Boizet-Bonhoure, T. Alatossava and P. Ritzenthaler. (1994). Evolution of the lysis modules in two related *Lactobacillus delbrueckii* bacteriophages. Microbiol. (submitted).
Foley, S., C. Daly, S. Bron and G.F. Fitzgerald. (1993). Incompatibility properties of the narrow-host-range lactococcal plasmid pCI305. FEMS Microbiol. Rev. **12**:Abst A14.
Gasson, M. J. and G.F. Fitzgerald. (1994). Gene transfer systems and transposition. In: Genetics and biotechnology of lactic acid bacteria (M. J. Gasson, W. M. De Vos. Eds.) Chapman and Hall, London. pp. 1-51.
Gasson, M.J. and W.M. de Vos (Eds.). (1994). Genetics and biotechnology of lactic acid bacteria. Chapman and Hall, London. pp. 1-310.
Godon, J.J., C.A. Shearman, K. Jury and M.J. Gasson. (1994). Cloning, sequencing and expression of the *Lactococcus lactis* sex factor aggregation gene *clu*A. Mol. Microbiol. **12**: 655-663.
Haandrikman, A.J., R. Meesters, H. Laan, W.N. Konings, J. Kok and G. Venema. (1991). Processing of the lactococcal extracellular serine proteinase. Appl. Environ. Microbiol. 57: 1899-1904.
Hayes, F., P. Vos, G. Fitzgerald, W.M. de Vos and C. Daly. (1991). Molecular organization of novel, narrow host-range, lactococcal plasmid pCI305. Plasmid **25**: 16-26.

Hertel, C., W. Ludwig, B. Pot, K. Kersters and K.H. Schleifer. (1993). Differentiation of lactobacilli occurring in fermented milk products by using oligonucleotide probes and electrophoretic protein profiles. Syst. Appl. Microbiol. **16**: 463-467.
I'Anson, K., H. Griffin, S Movahedi, F. Mulholland and M. J. Gasson. (1993). Cloning and analysis of peptidase gene(s) from *Lactococcus lactis*. FEMS Microbiol. Rev. **12**: P.74.
I'Anson, K., H. Griffin, S. Movahedi, F. Mulholland and M. J. Gasson. (1994). Cloning and sequencing of the glutamyl aminopeptidase gene, *pep* A, from *Lactococcus lactis*. (in preparation).
I'Anson, K., A.G. Griffin, S. Mohavedi, F. Mulholland and M.J. Gasson. (1994). Cloning and analysis of peptidase gene(s) from *Lactococcus lactis*. FEMS Microbiol. Rev. 12 P67.
Jiménez-Díaz, R., R.M. Rios Sánchez, M. Desmazeaud, J.L. Ruiz-Barba and J.-C. Piard. (1993). Plantaricins S and T, two new bacteriocins produced by *Lactobacillus plantarum* LPCO10 isolated from a green olive fermentation. Appl. Environ. Microbiol. **59**: 1416-1424.
Klaenhammer, T.R. and G.F. Fitzgerald. (1993). Bacteriophages and bacteriophage resistance. In: Genetics and biotechnology of lactic acid bacteria, (M. Gasson & W. De Vos Eds.) Chapman & Hall, London. pp. 106-168.
Kok, J. and W.M. de Vos. (1994). The proteolytic system of lactic acid bacteria. In: Genetics and biotechnology of lactic acid bacteria (M.J. Gasson and W.M. de Vos, Eds.) Chapman and Hall, PP. 169-210.
Kok, J., H. Holo, M.J. van Belkum, A.J. Haandrikman, and I.F. Nes. (1992). Non-Nisin Bacteriocins in Lactococci: Biochemistry, genetics, and mode of action. In: 'Bacteriocins of Lactic Acid Bacteria' (L. Steenson and D. Hoover, Eds.). American Society of Microbiology, Washington, D.C.
Kunji, E.R.S., E.J. Smid, R. Plapp, B. Poolman, and W.N. Konings. (1993). Di-tripeptides and oligopeptides are taken up via distinct transport mechanisms in *Lactococcus lactis*; J. Bacteriol. **175**: 2052-2059.
Laan, H., J. Kok, A.J. Haandrikman, G. Venema, and W.N. Konings. (1992). Localization and accessibility of antigenic sites of the extracellular serine proteinase of *Lactococcus lactis*. Eur. J. Biochem. **204**: 815-820.
Law, J., A. Haandrikman, G. Venema and K. Leenhouts. (1994). Random chromosomal insertions in *Lactococcus lactis*: identification of a mutation in the maltose metabolic pathway. J. Bacteriol. (submitted).
Law, J., A. Haandrikman, G. Venema and K. Leenhouts. (1994). Development of a general system for generating random chromosomal insertions. Appl. Environ. Microbiol. (submitted).
Law, J., P. Vos, F. Hayes, C. Daly, W.M. deVos and G.F. Fitzgerald. (1992). Cloning and partial sequencing of the proteinase gene complex from *Lactococcus lactis* subsp. *lactis* UC317. J. Gen. Microbiol. **138**: 709-718.
LeLoir, Y., A. Gruss, S.D. Ehrlich, P. Langella. (1994). Direct screening of recombinants in Gram-positive bacteria using the secreted staphylococcal nuclease as reporter. J. Bacteriol. (in press).
Mayo, B., J. Kok, K. Venema, W. Bockelmann, M. Teuber, H. Reinke and G. Venema. (1991). Molecular cloning and sequence analysis of the X-prolyl-dipeptidyl aminopeptidase gene from *Lactococcus lactis* subsp. *cremoris*. Appl. Environ. Microbiol. **57**: 38-44.
Mayo, B., J. Kok, W. Bockelmann, A. Haandrikman, K.J. Leenhouts and G. Venema. (1993). The effect of X-prolyl-dipeptidyl-aminopeptidase deficiency in *Lactococcus*. Appl. Environ. Microbiol. **59**: 2049-2055.
Mierau, I., A.J. Haandrikman, O. Velterop, P.S.T. Tan, K.J. Leenhouts, W.N. Konings, G. Venema, and J. Kok. (1994). Tripeptidase gene (*pepT*) of *Lactococcus lactis*: molecular cloning and nucleotide sequence of *pep*T and construction of a chromosomal deletion mutant. J. Bacteriol. **176**: 2854-2861.
Mierau, I., P.S.T.Tan, A.J. Haandrikman, J. Kok, K.J. Leenhouts, W.N. Konings, and G. Venema. (1993). Cloning and sequencing of the gene for a lactococcal endopeptidase, an enzyme with sequence similarity to mammalian enkephalinase. J. Bacteriol. **175**: 2087-2096.
Neubauer, H., E. Glaasker, W.P. Hammes, B. Poolman and W. N. Konings. (1994).Mechanism of maltose uptake and glucose excretion in *Lactobacillus sanfrancisco*. J. Bacteriol. (in press).
Parreira, R., R. Valyasevi, S.D. Ehrlich and M.C. Chopin. (1994). *abi-416* (abiB) encoded abortive infection mechanism induces degradation of lactococcal phage RNAs. (in preparation).

Patarata, L., M. Serpa-Pimentel, B. Pot, K. Kersters and A. Mendes-Faia. (1994). Identification of Lactic acid bacteria isolated from Portuguese wines and musts by SDS-PAGE. J. Appl. Microbiol. **76**: 288-293.
Piard, J.C., O.P. Kuipers, H.S. Rollema, M. Dezmazeaud and W.M. de Vos. (1993). Structure, function, and expression of the *lct* gene for lacticin 481, a novel lantibiotic produced by *Lactococcus lactis*. J. Biol. Chem. **268**: 1636-1638.
Pot, B., W. Ludwig, K. Kersters and K.H. Schleifer. (1994). Taxonomy of lactic acid bacteria. In: Bacteriocins of lactic acid bacteria: microbiology, genetics and application. L. de Vuyst and E.J. Vandamme Eds. Elsevier Science Publishers Ltd. Amsterdam. pp. 14-90.
Pot, B., C. Hertel, W. Ludwig, P. Descheemaeker, K. Kersters, and K.H. Schleifer. (1992). Identification and classification of *Lactobacillus acidophilus, L. gasseri* and *L. johnsonii* strains by SDS-PAGE and rRNA-targeted oligonucleotide probe hybridization. J. Gen Microbiol. **139**: 513-517.
Ramos, A., K.N. Jordan, T.M. Cogan and H. Santos. (1994). Nuclear magnetic resonance studies of citrate and glucose co-metabolism by *Lactococcus lactis*. Appl. Environ. Microbiol. **60**: (in press).
Ramos, A., K.N. Jordan, T.M. Cogan and H. Santos. (1994). ^{13}C-NMR studies of citrate and glucose co-metabolism by *Lactococcus lactis*. Appl. Environ. Microbiol. **60**: 1739-48.
Ramos, A., B. Poolman, H. Santos, J.S. Lolkema and W.N. Konings. (1994). Uniport of anionic citrate and proton consumption in citrate metabolism generate a proton motive force in *Leuconostoc oenos*. J. Bacteriol. (in press).
Riguez, J.M., L.M. Cintas, P. Casaus, N. Horn, H. Dodd, P.E. Hernandez and M.J. Gasson. (1994). Isolation of nisin-producing *Lactococcus lactis* strains from dry fermented sausages. J. Appl. Bacteriol. (submitted).
Ruiz-Barba, J.L., D.P. Cathcart, P.J. Warner and R. Jimenez-Diaz. (1993). Use of *Lactobacillus plantarum* LPCO-10, a bacteriocin producer, as a starter culture in Spanish style green olive fermentations. FEMS Microbiol. Rev. **12**:128.
Ruiz-Barba, J.L., D.P. Cathcart, P.J. Warner and R. Jiménez-Díaz. (1994). Use of *Lactobacillus plantarum* LPCO10, a bacteriocin producer, as a starter culture in Spanish-style green olive fermentations. Appl. Environ. Microbiol. **60**: (in press).
Ruiz-Barba, J.L., D.P. Cathcart, R. Jiménez-Díaz and P.J. Warner. (1993). Purification and characterisation of plantaricin S, a bacteriocin produced by *Lactobacillus plantarum* LPCO10. FEMS Microbiol. Rev. **12**:125.
Ruiz-Barba, J.L., D.P. Cathcart, H. Holo, I. Nes, K.H. Sletten, R. Jiménez-Díaz and P.J. Warner. (1994). Purification and characterisation of *plantaricin S*, a bacteriocin produced by *Lactobacillus plantarum* LPCO10. Appl. Environ. Microbiol.(in preparation).
Schleifer, K.H., M. Ehrmann, U. Krusch and H. Neve. (1991). Revival of the species *Streptococcus thermophilus* (ex Orla-Jensen, 1919) nom. rev. Syst. Appl. Microbiol. **14**: 386-388.
Shearman, C. A., S. Hertwig, M. Teuber and M.J. Gasson. (1991). Characterisation of the prolate-headed lactococcal bacteriophage ΦvML3: location of the lysin gene and its DNA homology with other prolate-headed phages. J. Gen. Microbiol. **137**:1285-1291.
Stoddard, G.W., J.P. Petzel, M.J. van Belkum, J. Kok, and L.L. McKay. (1992). Molecular analyses of the Lactococcin A gene cluster from *Lactococcus lactis* subsp. *lactis* biovar *diacetylactis* WM4. Appl. Environ. Microbiol. **58**: 1952-1961.
Tan, P.S.T., I.J. van Alen-Boerrigter, B. Poolman, R.J. Siezen, W.M. de Vos and W.N. Konings. (1992). Characterisation of the *Lactococcus lactis pepN* gene encoding an aminopeptidase homologous to mammalian aminopeptidase N. FEBS Lett. **306**: 9-16.
Thompson, J.K., K.J. McConville and M.A. Collins. (1993). Variability of plasmid DNA and associated phenotypes in *Lactobacillus helveticus* DSM20075. Abstract FEMS Microbiology Reviews. 4th Symposium on Lactic Acid Bacteria Genetics, Metabolism. and Applications, Noordwijkerhout, The Netherlands, 5-9 Sept 1993.
Tsakalidou, E., E. Manolopoulou, E. Kabaraki, E. Zoidou, B. Pot, K. Kersters, and G. Kalantzopoulos. (1994). The combined use of whole-cell protein extracts for the identification (SDS-PAGE) and enzyme activity screening of lactic acid bacteria isolated from traditional Greek dairy products. Syst. Appl. Microbiol. (in press).
Tynkkynen, S., G. Buist, E. Kunji, J. Kok, B. Poolman, G. Venema and A. Haandrikman. (1993). Genetic and biochemical characterisation of the oligopeptide transport system of *Lactococcus lactis*. J. Bacteriol. **175**:7523-7532.

Van Belkum, M.J., J. Kok, G. Venema, H. Holo, I.F. Nes, W.N. Konings and T. Abee. (1991). The bacteriocin lactococcin A specifically increases permeability of lactococcal cytoplasmic membranes in a voltage-independent, protein-mediated manner. J. Bacteriol. 173:7934-7941.
Van Belkum, M.J., B.J. Hayema, J. Kok, G. Venema, H. Holo, I.F. Nes, W.N. Konings, and T. Abee. (1992). Lactococcal bacteriocins: genetics and mode of action. In: Bacteriocins, microcins and lantibiotics (R. James, C. Lazdunski and F. Pattus, Eds.) Springer-Verlag, Berlin. pp. 59-69.
Van de Berg, D.J.C., A.Smits, B. Pot, A.M. Ledeboer, K. Kersters, J.M.A. Verbakel and C.T. Verrips. (1993). Isolation, screening and identification of lactic acid bacteria from traditional food fermentation processes and culture collections. Food Biotech. 7:189-205.
Van der Lelie, D., F. Chavarri, G. Venema, M.J. Gasson. (1991). Identification of a new genetic determinant for cell aggregation associated with lactose plasmid transfer in *Lactococcus lactis*. Appl. Environ. Microbiol. **57**: 201-206.
Van Rooijen, R. J., M.J. Gasson and W.M. de Vos. (1992). Characterisation of the *Lactococcus lactis* lactose operon promoter: contribution of flanking sequences and LacR repressor to promoter activity. J. Bacteriol. **174**: 2273-2280.
Vasala, A., L. Dupont, M. Baumann, P. Ritzenthaler and T. Alatossava. (1993). Molecular comparison of gene clusters encoding the structural protein of two related *Lactobacillus delbrueckii* bacteriophages. J. Virol. 67: 3061-3068.
Venema, K., T. Abee, A.J. Haandrikman, K.J. Leenhouts, J. Kok, W.N. Konings, and G. Venema. (1993). Mode of action of lactococcin B, a thiol-activated bacteriocin from *Lactococcus lactis*. Appl. Environ. Microbiol. **59**: 1041-1048.
Vogel, R. F., G. Bocker, P. Stolz, M. Ehrmann, D. Fanta, W. Ludwig, B. Pot, K. Kersters, K.H. Schleifer and W. P. Hammes. (1994). Identification of lactobacilli from sourdough and description of *Lactobacillus pontis* sp. nov. Int. J. System. Bacteriol. **44**: 223-229.
Vos, P, I.J. Boerrigter, G. Buist, A.J. Haandrikman, M. Nijhuis, M.B. de Reuver, R.J. Siezen, G. Venema, W.M. de Vos and J. Kok. (1991). Engineering of the *Lactococcus lactis* proteinase by construction of hybrid enzymes. Protein Eng. **4**: 479-484.
Wells, J. M., P.W. Wilson, P.M. Norton, M.J. Gasson, M. J. and R.W.F. Le Page. (1993). The use of T7 RNA polymerase for regulated heterologous gene expression in *Lactococcus lactis*. Mol. Microbiol. **8**: 1155-1162.

Individual publications.

Anderson, S.V. and M.A. Collins. (1993). The Characterisation of the α-galactosidase gene expression and stability in relation to lactose transport in *Lactobacillus helveticus*. Abstract FEMS Microbiology Reviews. 4th Symposium on Lactic Acid Bacteria Genetics, Metabolism & Applications, Noordwijkerhout, The Netherlands, 5-9 September 1993.
Anderson, S.V. and M.A. Collins. (1994). Factors affecting the genetic basis underlying the occurrence of spontaneous *lac*$^-$ mutants in *Lactobacillus helveticus* (in press).
Arendt, E.K., M. van de Guchte, A. Coffey, C. Daly and G.F. Fitzgerald. (1993). Molecular genetics of bacteriophages of lactic acid bacteria. Le Lait **73**: 191-198.
Arendt, E.K., C. Daly, G.F. Fitzgerald and M. van de Guchte. (1994). Molecular characterisation of the lactococcal bacteriophage Tuc2009 and identification and analysis of genes encoding lysin, a putative holin, and two structural proteins. Appl. Environ. Microbiol. **60**: (in press).
Bardowski, J., S.D. Ehrlich and A. Chopin. (1994). BglR protein, which belongs to the BglG/SacY family of transcriptional antiterminators, is involved in β-glucoside utilization in *Lactococcus lactis*. J. Bacteriol. (in press).
Bardowski, J., S.D. Ehrlich, A. Chopin. (1992). Genes for tryptophan biosynthesis in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. 174: 6563-6570.
Bardowski, J., S.D. Ehrlich, A. Chopin. (1994). An RNA-binding transcriptional antiterminator regulates β-glucoside assimilation in *Lactococcus lactis* subsp. *lactis*. In: 'Genetics of streptococci, enterococci and lactococci', (J.J. Ferretti Ed.) (in press).
Bardowski, J., S. D. Ehrlich, and A. Chopin. (1992). Tryptophan biosynthesis genes in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol., **174**: 6563-6570.
Beimfohr, C., A. Krause, R. Amann, W. Ludwig and K.H. Schleifer. (1993). *In situ* identification of lactococci, enterococci and streptococci. Syst. Appl. Microbiol. **16**: 450-456.
Biswas, I., A. Gruss, S.D. Ehrlich and E. Maguin. (1993). High efficiency gene inactivation and replacement system for Gram positive bacteria. J. Bacteriol. **175**: 3628-3635.

Bobillo, M. and V.M. Marshall. (1991). Effect of salt and culture aeration on lactate and acetate production by *Lactobacillus plantarum*. Food Microbiol. **8**: 153-160.
Bockelmann, W., T. Hoppe-Seyler, and K. Heller. (1994). Purification and characterisation of an endopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*. (submitted).
Bockelmann, W., H.-P. Beuck, S. Lick and K. Heller. (1994). Purification and characterisation of a new tripeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus* B14. (submitted).
Bockelmann, W., M. Fobker and M. Teuber. (1991). Purification and characterisation of the X-prolyl-dipeptidyl-aminopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Lactobacillus acidophilus*. Int. Dairy Journal **1**: 51-66.
Bockelmann, W., Y. Schulz and M. Teuber. (1992). Purification and characterisation of an aminopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*. Int. Dairy Journal **2**: 95-107.
Boizet-Bonhoure, B., L. Dupont, M. Mata and P. Ritzenthaler. (1993). Site-specific integration region of the bacteriophage mv4 of *Lactobacillus delbrueckii* subsp. *bulgaricus*. FEMS Microbiol Rev **12**: 22.
Brink, B. ten., M. Minekus, J.M.B.M. van de Vossen, R.J. Leer and J.H.J. Huis in't Veld. (1994). Antimicrobial activity of lactobacilli: preliminary characterisation and optimization of production of acidocin b, a novel bacteriocin produced by *Lactobacillus acidophilus* M46. J. Appl. Bacteriol. (in press).
Brink B. ten and J.H.J. Huis in't Veld. (1994). Comparison of the antimicrobial spectrum of bacteriocins produced by lactic acid bacteria studied in the EC-BRIDGE Programme. Food Microbiol. (submitted).
Bruinenberg, P.G., P. Doesburg, A.C. Alting, F.A. Exterkate, W.M. de Vos and R.J. Siezen. (1994). Evidence for a large dispensible segment in the subtilisin-like catalytic domain of the *Lactococcus lactis* cell envelope proteinase. Protein Eng. (in press).
Bruinenberg, P.G., P. Vos, and W.M. de Vos. (1992). Proteinase overproduction in *Lactococcus lactis*: Regulation and effect on growth and acidification in milk. Appl. Environ. Microbiol. **58**: 78-84.
Bruinenberg, P.G., P. Vos, F.A. Exterkate, A.C. Alting, W.M. de Vos and R.J. Siezen. (1993). Engineering stability and specificity of the *Lactococcus lactis* SK11 proteinase. In: Stability and stabilization of enzymes (W. van der Tweel, A. Harder and R.M. Buitelaar, Eds.) Elsevier Science Publishers. pp. 231-238.
Casey, C.N. (1991). Molecular and physical characterisation of lysogenic and lytic lactococcal bacteriophages. PhD. Thesis, National University of Ireland, Cork, IE.
Casey, C.N., E. Morgan, C. Daly and G.F. Fitzgerald. (1993). Characterisation and classification of virulent lactococcal bacteriophages isolated from a Cheddar cheese plant. J. Appl. Bacteriol. **74**: 268-275.
Casey, J., C. Daly and G.F. Fitzgerald. (1992). Controlled integration into the *Lactococcus* chromosome of the pCI829-encoded abortive infection gene from *Lactococcus lactis* ssp. *lactis* UC811. Appl. Environ. Microbiol. **58**: 3283-3291.
Casey, J., C. Daly and G.F. Fitzgerald. (1991). Chromosomal integration of plasmid DNA by homologous recombination in *Enterococcus faecalis* and *Lactococcus lactis* subsp. *lactis* hosts harbouring Tn919. Appl. Environ. Microbiol. **57**: 2677-2682.
Casey, J. (1992). Chromosomal integration strategies for *Lactococcus* and *Enterococcus*: development of Tn919-based integration vectors. PhD. Thesis, National University of Ireland, Cork, IE.
Caso J.L. and J. E. Suárez. (1994). Detection and functional analysis of *Lactobacillus* endodeoxyribonucleases. FEMS Microbiol. Lett. (in preparation).
Caso, J. L., C. G. Reyes, M. Herrero, A. Montilla, A. Rodríguez and J. E. Suárez. (1994). Isolation and characterisation of temperate and virulent bacteriophages of *Lactobacillus plantarum*. J. Dairy Sci. (submitted).
Chapot-Chartier, M.P., M. Nardi, M.C. Chopin, A. Chopin and J.C. Gripon. (1993). Cloning and sequencing of *pepC*, a cysteine aminopeptidase gene from *Lactococcus lactis* subsp. *cremoris AM2*. Appl. Environ. Microbiol. **59**: 330-333.
Chopin, A. (1993). Organization and regulation of genes for amino acid biosynthesis in lactic acid bacteria. FEMS Microbiol. Rev. **12**: 21-38.
Chopin, A., J. Bardowski, R. Raya and S.D. Ehrlich. (1993). Genetic regulation in *Lactococcus lactis*: the model of tryptophan biosynthesis. Le Lait. **73**: 119-126.
Cintas, L.M., W.L. Moreira, J.M. Rodriguez, O.J. Sobrino, M.F. Fernandez, B. Sanz and P.E. Hernandez. (1992). Antimicrobial activity of lactic acid bacteria from meat origin against

selected indicator microorganisms. Proceedings of the 38th International congress of meat science and technology, Clermont-Ferrand, France pp. 643-646.
Cocconcelli P.S., B. Riboli, M. Elli, V. Bottazzi and L. Morelli. (1994). Genetic analysis of the replication region of the *Lactobacillus* plasmid vector pPSC 22. (submitted).
Coffey, A.G., A.J. Harrington, K. Kearney, C. Daly and G.F. Fitzgerald. (1994). Nucleotide sequence and structural organization of the small, broad-host-range plasmid pCI411 from *Leuconostoc lactis* 533. Microbiol. (in press).
Coffey, A.G., G.F. Fitzgerald and C. Daly. (1991). Cloning and characterisation of determinant for abortive infection of bacteriophage from lactococcoal plasmid pCI829. J. Gen. Microbiol. **137**: 1355-1362.
Coffey, A., V. Costello, C. Daly and G.F. Fitzgerald. (1991). Plasmid-encoded bacteriophage insensitivity in members of the genus *Lactococcus* with special reference to pC1829. In: Genetics and molecular biology of streptococci, lactococci and enterococci. Ed. (G.M. Dunney et. al. Ed.) American Society for Microbiology. pp. 131-135.
Collins, M.A. and J.K. Thompson. (1992). Lactose metabolism and lactose gene sequence homologies amongst lactobacilli. J. Appl. Bact. **73**: 237-242.
Collins, M.A., S.V. Anderson and J.K. Thompson. (1993). The location and stability of *lac* genes in *Lactobacillus helveticus*. Abstract 6th European congress on biotechnology, Firenze, Italy 13-17 June 1993. Vol. 1.
Couto, J.A., and T. Hogg. (1994). Diversity of ethanol tolerant lactobacilli isolated from Douro fortified wine. J. Appl. Bacteriol. **76**: 487-491.
Cuesta, E. P., J. E. Suárez and A. Rodríguez. (1994). Incidence of lysogeny in wild lactococcal strains. J. Dairy Sci. (submitted).
Delorme, C., J.J. Godon, S.D. Ehrlich and P. Renault. (1993). Gene inactivation in *Lactococcus lactis* II. Histidine biosynthesis. J. Bacteriol. **175**: 4391-4399.
Delorme, C., S. D. Ehrlich, and P. Renault. (1992). Histidine biosynthesis genes in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. **174**: 6571-6579.
Delorme, C., J.J. Godon, S.D. Ehrlich, P. Renault. (1994). Mosaic structure of large regions of the *Lactococcus lactis* subsp. *cremoris* chromosome. Mol. Microbiol. (in press).
David, S., H. Stevens, M. van Riel, G. Simons and W.M. de Vos. (1992). *Leuconostoc lactis* β-galactosidase is encoded by two overlapping genes. J. Bacteriol. **174**: 4475-4481.
Davis, R. (1992). Development of *in vitro* gene transfer in *Lactobacillus* species and restriction-modification in Lactococcus: biochemical and genetic analysis of the *Scr*FI system. PhD. Thesis, National University of Ireland, Cork, IE.
Dellaglio, F., H. de Roissart, S. Torriani, M.C. Curk and D. Janssens. (1994). Caracteristiques generales des bacteries lactiques. In: Bacteries Lactiques (I.H. de Roissart and F.M. Luquet Eds.) pp. 25-116. Lorica, Uriage, France.
De Revel, G., A.B. Capela and T. Hogg. (1994). A pre-spoilage marker for bacterial activity in fortified wine, conversion of L-malic to L-lactic acid. Letts. Appl. Microbiol. (in press).
De Vos, W.M., I. Boerrigter, P. Vos, P. Bruinenberg and R.J. Siezen. (1991). Proteinase production, processing and engineering in *Lactococcus lactis* SK11. In: Genetics and Molecular Biology of streptococci, lactococci and enterococci (G.M. Dunny, P.P.Cleary and L.L. McKay, Eds.), American Society for Microbiology, Washington, DC. pp. 115-119.
De Vos, W.M. (1992). Engineering of Lactic Acid Bacteria for Improved Production of Foods, Flavors and Additives. In: Harnessing Biotechnology in the 21st Century. (R. Lodisch & A. Bose, Eds.) American Chemical Society. pp. 524-527.
De Vos, W.M. (1993). Future prospects for research and application of nisin and other bacteriocins. In: Bacteriocins of lactic acid bacteria (D. Hoover and L.Steenson, Eds.) Academic Press. pp. 249-265.
De Vos, W.M. (1993). Engineering of lactic acid bacteria for improved food fermentations. In: Developing agricultural biotechnology in the Netherlands (D.H. Vuijk, J.J. Dekkers, and H.C. van der Plas, Eds.) Pudoc, Wageningen. pp. 231-235.
De Vos, W.M. and G. Simons. (1994). Gene cloning and expression systems in lactococci. In: Genetics and biotechnology of lactic acid bacteria (M.J. Gasson and W.M. de Vos Eds.) Chapman and Hall. pp. 52-105.
De Vos, W.M. and R.J. Siezen. (1994). Engineering pivotal proteins for Lactococcal proteolysis. Society for Protein Chemistry. Reading, UK (in press).

Devriese, L.A., B. Pot, L. van Damme, K. Kersters, and F. Haesebrouck. (1994). Identification of *Enterococcus* species isolated from food products of animal origin. J. Appl. Bacteriol. (in press).
Devriese, L. A. and B. Pot. (1994). The genus *Enterococcus*. In: The lactic acid bacteria: Volume 2: The genera of lactic acid bacteria. (W. Holzapfel and J. Wood, Eds.) (in press).
Dodd, H. M., N. Horn, Z. Hao and M.J. Gasson. (1992). A lactococcal expression system for engineered nisins. Appl. Environ. Microbiol. **58**: 3683-3693.
Dodd, H. M., N. Horn, S. Swindell and M.J. Gasson. (1992). Tn5301, a lactococcal transposon encoding genes for nisin biosynthesis. In: NATO ASI Series Vol. H 65 bacteriocins, microcins and lantibiotics (R. James, C. Lazdunski, F. Pattus, Eds.). Springer-Verlag Berlin Heidelberg.
Dodd, H. and M.J. Gasson. (1994). Bacteriocins of lactic acid bacteria. In: Applied genetics of lactic acid bacteria (M.J. Gasson, W.M. De Vos, Eds.). Blackie, Edinburgh. pp. 211-245.
Dodd, H.M., N. Horn and M.J. Gasson. (1994). Characterisation of IS905, a new multicopy insertion sequence identified in lactococci. J. Bacteriol. **176**: 3393-3396.
Dodd, H.M., N. Horn and M.J. Gasson. (1993). Production of novel nisins. FEMS Microbiol. Rev. **12**: 122.
Dossmann, M., R. F. Vogel and W. P. Hammes. (1994). A CSTR system for the investigation of meat lactobacilli. (in preparation).
Dossmann, M., R.F. Vogel and W. P. Hammes. (1994). Behaviour of *Lactobacillus curvatus* and *L. sake* in continuous cultures simulating the conditions of sausage fermentation. (in preparation).
Duckworth, A.W. , J.L. Ruiz-Barba and P.J. Warner. (1993). Construction of integrating derivatives of pHV60 for use in *Lactobacillus plantarum*. FEMS Microbiol. Lett. **114**: 267-272.
Dupont, L. (1994). Organization, expression and evolution of the genes of the temperate bacteriophage mv4 of *Lactobacillus delbrueckii* subsp. *bulgaricus*. PhD. Thesis., Universite Paul Sabatler, Toulouse, FR.
Duwat, P., S.D. Ehrlich and A. Gruss. (1992). PCR cloning and sequencing of the *Lactococcus lactis* subsp. *lactis recA* gene. Appl. Env. Microbiol. **58**: 2674-2678.
Duwat, P., S.D. Ehrlich and A. Gruss. (1992). A general method to clone recA genes of Gram-positive bacteria by polymerase chain reaction. J. Bacteriol. **174**: 5171-5175.
Eaton, T. J., C.A. Shearman and M.J. Gasson. (1993). The use of bacterial luciferase genes as reporter genes in *Lactococcus*: regulation of the *Lactococcus lactis* subsp. *lactis* lactose operon. J. Gen. Microbiol. **139**: 1495-1501.
Eaton, T. J., C.A. Shearman and M.J. Gasson. (1993). Cloning and sequence analysis of the *dnaK* gene region of *Lactococcus lactis* subsp. *lactis*. J. Gen. Microbiol. **139**: 3253-3264.
Ehrmann, M., W. Ludwig, and K.H. Schleifer. (1992). Species specific oligonucleotide probe for the identification of *Streptococcus thermophilus*. Syst. Appl. Microbiol. **15**: 453-455.
Ehrmann, M., W. Ludwig and K.H. Schleifer. (1994). Reverse dot blot hybridization: a useful method for the direct identification of lactic acid bacteria in fermented food. FEMS Microbiol. Lett. **117**: 143-149.
Exterkate, F.A., A.C. Alting and P. Bruinenberg. (1993). Diversity of cell-envelope proteinase specificity among strains of *Lactococcus lactis* and the relation with charge characteristics of the substrate binding region. Appl. Environ. Microbiol. **59**: 3640-3647.
Firme, M.P., M.C. Leitao and M.V. San-Romao. (1994). The metabolism of sugar and malic acid by *Leuconostoc oenos*: effect of malic acid, pH and aeration conditions. J. Appl. Bacteriol. **76**: 173-181.
Fitzgerald, G.F., A.G. Coffey and C. Daly. (1993). Molecular manipulations of *Lactococcus* starter cultures for food fermentations. J. Chem. Technol. Biotech. **58**: 195-199.
Foley, S., C. Daly and G.F. Fitzgerald. (1993). Incompatibility properties of the narrow-host-range lactococcal plasmid pCI305. J. Gen. Microbiol. **139**: 1271-1276.
Garvey, P., G.F. Fitzgerald and C. Hill. (1993). The conjugal lactococcal plasmid pNP40 encodes multiple bacteriophage resistance mechanisms. FEMS Microbiol. Rev. **12**:Abst F14.
Gasson, M.J. (1993). Progress and potential in the biotechnology of lactic acid bacteria. FEMS Microbiol. Rev. 12: 3-20.
Gasson, M. J., S. Swindell, S. Maeda and H.M. Dodd. (1992). Molecular rearrangement of lactose plasmid DNA associated with high frequency transfer and cell aggregation in *Lactococcus lactis* 712. Mol. Microbiol. **6**: 3213-3223.
Geis, A., T. Janzen, M. Teuber and F. Wirsching. (1992). Mechanism of plasmid mediated bacteriophage resistance in lactococci. FEMS Microbiol. Lett. **94**: 7-14.

Godon, J .J., M. C. Chopin, and S. D. Ehrlich. (1992). Branched-chain amino acid biosynthesis genes in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. **174**: 6580-6589.
Godon, J. J., C. Delorme, S. D. Ehrlich, and P. Renault. (1993). Divergence of genomic sequences between *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*. Appl. Environ. Microbiol. **58**: 4045-4047.
Godon, J.J., C. Delorme, J. Bardowski, M. C. Chopin, S. D. Ehrlich, and P. Renault. (1993). Gene inactivation in *Lactococcus lactis* I. Branched-chain amino acid biosynthesis. J. Bacteriol. **175**: 4383-4390. .
González, B., P. Arca, B. Mayo and J. E. Suárez. (1994). Detection, purification and partial characterisation of plantaricin C, a bacteriocin produced by a *Lactobacillus plantarum* strain of dairy origin. Appl. Environ. Microbiol. (in press).
Grenov, B. Fermentation studies of *P. acidilactici* PA-1, producing pediocin. Chr. Hansens TR00648.
Grenov, B. Fermentation studies of *L. bavaricus* MI401, producing bavaricin. Chr. Hansens TR00649.
Griffin, H. G., S.R. Swindell and M.J. Gasson. (1992). Cloning and DNA sequence analysis of the genes for L-lactate dehydrogenase from *Lactococcus lactis*: evolutionary relationships between 21 different *Idh* genes. Gene **122**: 193-197.
Griffin, H. G. and M.J. Gasson. (1993). The regulation of expression of the *Lactococcus lactis* lactose operon. Lett. Appl. Microbiol. **17**: 92-96.
Griffin, H.G., K.J. I'Anson and M.J. Gasson. (1993). Rapid isolation of genes from bacterial lambda libraries by direct polymerase chain reaction screening. FEMS Microbiol. Lett. **112**: 49-54.
Griffin, H., K. Benson, S. Swindell, M.J. Gasson. (1993). Analysis and exploitation of lactococcal genes relevant to diacetyl production. FEMS Microbiol. Rev. **12**: 22.
Haandrikman, A.J., J. Kok and G. Venema. (1991). Lactococcal proteinase maturation protein PrtM is a lipoprotein. J. Bacteriol. **173**: 4517-4525.
Hagling, A., E.R.S. Kunji, K.J. Leenhouts, B. Poolman and W.M. Konings. (1994).
The di- and tripeptide transport protein of *Lactococcus lactis*. A new type of bacterial peptide transporter. J. Biol. Chem. **269**: 11391-11399.
Harrington, A. (1994). Genetic Analysis of *Leuconostoc* species; a study of citrate permease activity and plasmid replication. Ph.D. Thesis, National University of Ireland, Cork, IE.
Hayes, F., C. Daly and G.F. Fitzgerald. (1993). Physical analysis of pCI305 fusion plasmids in *Lactococcus lactis* subsp. *lactis*. FEMS Microbiol. Lett. **106**: 295-300.
Hernandez, P.E., J.M. Rodriguez, L.M., Cintas, W.L. Moreira, O.J. Sobrino, M.F. Fernandez, B. Sanz and P.E. Hernandez. (1994). Inhibition of *Listeria monocytogenes* by *Lactobacillus* sake strains of meat origin. Meat Science (in press).
Herrero, M., C. G. Reyes, J. L. Caso and J. E. Suárez. (1994). Characterizstion of Φ393-A2, a bacteriophage that infects *Lactobacillus casei*. Microbiol. (in press).
Hill, C., C. Daly, and G.F. Fitzgerald. (1991). Isolation of chromosomal mutations of *Lactococcus* subsp. *lactis* biovar. *diacetylactis* 18-16 after introduction of Tn919. FEMS Microbiol Lett. **81**: 135-140.
Horn, N., S. Swindell, H.M. Dodd, and M.J. Gasson. (1991). Nisin biosynthesis genes are encoded by a novel conjugative transposon. Mol. Gen. Genet. **228**: 129-135 .
Jelle, B. Simulated meat trials with *P. acidilactici* PA-1 and *L. bavaricus* MI401 and effect against *L. monocytogenes*. Chr. Hansens TR00650.
Jelle, B. Antimicrobial activity of *P. acidilactici* PA-1 and *L. bavaricus* MI401 towards *L. monocytogenes* in fermented sausages. Chr. Hansens TR00651.
Juillard, V., H. Laan, E.R.S. Kunji, C.M. Jeronimum-Stratingh, A.P. Bruins and W.M. Konings. (1994). Hydrolysis of bovine β-casein by purified proteinase of *Lactococcus lactis* subsp. *cremoris* Wg2 (P_I-type). Appl. Environ. Microbiol. (submitted).
Kersters, K., B. Pot, D. Dewettinck, U. Torck, M. Vancanneyt, L. Vauterin and P. Vandamme. (1994). Identification and typing of bacteria by protein electrophoresis. In: Identification of bacteria: current status, future prospects. (F.G. Priest, A. Ramos-Cormenzana and B. Tindall Eds.) Plenum Publishing Corporation (in press).
Kersters, K. and B. Pot. (1992). Classification and identification methods for lactic acid bacteria, with emphasis on protein gel electrophoresis. In: Les Bacteries Lactiques (lactic acid bacteria). Actes du Colloque Lactic'91, Caen pp. 33-40 Adria Normandie — Centre de publications de l'Université de Caen, Caen, France.

Kiewiet, R., J. Kok, J.F.M.L. Seegers, G. Venema, and S. Bron. (1993). The mode of replication is a major factor in segregational plasmid instability in *Lactococcus lactis*. Appl. Environ. Microbiol. **59**: 358-364.
Kiewiet, R., S. Bron, K. de Jonge, G. Venema, and J.F.M.L. Seegers. (1993). Theta replication of the lactococcal plasmid pWV02 . Mol. Microbiol. **10**: 319-327.
Klein, J.R., C. Ulrich, and R. Plapp. (1993). Characterisation and sequence analysis of a small cryptic plasmid from *Lactobacillus curvatus* LTH683 and its use for construction of new *Lactobacillus cloning* vectors. Plasmid **30**: 14-29.
Klein, J.R., U. Schmidt and R. Plapp. (1994). Cloning, heterologous expression, and sequencing of the proline iminopeptidase gene, pepI, from *Lactobacillus delbrückii* subsp. *lactis* DSM 7290. Microbiol. 140: 1133-1139.
Klein, J.R., U. Klein, M. Schad, and R. Plapp. (1993). Cloning, DNA sequence analysis and partial characterisation of pepN, a Lys-amino-peptidase from *Lactobacillus delbrückii* ssp. *lactis* DSM7290. Eur. J. Biochem. **217**: 105-114.
Kojic, M., M.J. Gasson and L. Topisirovic. (1993). Characterisation of large conjugative plasmid of *Lactococcus lactis* subsp. *lactis* biovar *diacetylactis*. FEMS Microbiol. Rev. **12**: 20.
Kok, J. (1991). Special purpose vectors for lactococci. In: Genetics and molecular biology of streptococci, lactococci, and enterococci (G. Dunny, P. Cleary and L. McKay Eds.) ASM, Washington, pp. 97-102.
Kok, J. (1991). Proteinase genes of cheese starter cultures, Biochem. Soc. Trans. **19**: 670-674.
Kok, J. (1993). Genetics of proteolytic enzymes of lactococci and their role in cheese flavor development. J. Dairy Sci. **76**: 2056-2064.
Kuipers, O.P., H.J. Boot and W.M. de Vos. (1991). Improved Site-directed Mutagenesis Method Using PCR. 1991. Nucl. Acids Res. **19**: 4558.
Kunji, E.R.S., A. Hagting, C.J. de Vries, V. Juillard, A. Haandrikman, B. Poolman and W.N. Konings. (1994). Central role of the oligopeptide transport system in the proteolytic pathway of *Lactococcus lactis*: transport of β-casein derived peptides. J. Biol Chem. (submitted).
Kyriadkidis, S., G. Sakellaris and T. Sotiroudis. (1993). Is protein phosphorylation a control mechanism for the degradation of caseins by lactic acid bacteria? The detection of an extracellular acid phosphatase activity. Lett. Appl. Microbiol. **16**: 295-298.
Laan. H., H. Bolhuis, B. Poolman, T. Abee, and W.N. Konings. (1993). Regulation of proteinase synthesis in *Lactococcus lactis*. Acta Biotech. **13**: 94-101.
Lahbib-Mansais, Y., B. Boizet, L. Dupont, M. Mata and P. Ritzenthaler. (1992). Characterisation of a temperate bacteriophage of *Lactobacillus delbrueckii* subsp. *bulgaricus* and its interactions with the host cell chromosome. J. Gen. Microbiol. **138**: 1139-1146.
Langella, Ph., S.D. Ehrlich and A. Gruss. (1993). Efficient plasmid mobilization by pIP501 in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. **174**: 5806-5813.
Law, J., G.F. Fitzgerald, T. Uniacke, C. Daly and P.F. Fox. (1993). The contribution of lactococcal starter proteinases to proteolysis in cheddar cheese. J. Dairy. Sci. **76**: 2455-2467.
Law, J., G.F. Fitzgerald, C. Daly, P.F. Fox and N.Y. Farkye. (1992). Proteolysis and flavour development in Cheddar cheese made with the single starter strains *Lactococcus lactis* subsp. *lactis* UC317 and *Lactococcus lactis* subsp. *cremoris* HP. J. Dairy. Sci. **75**: 1173-1185.
Law, J. (1991). Biochemistry and Genetics of the proteolytic system in *Lactococcus*. PhD. Thesis, National University of Ireland, Cork, IE.
Leenhouts, K.J., J. Kok and G. Venema. (1991). Lactococcal plasmid pWVO1 as an integration vector for Lactococci. Appl. Environ. Microbiol. **57**: 2562-2567.
Leenhouts, K.J., B. Tolner, S. Bron, J. Kok, G. Venema and J.F.M.L. Seegers. (1991). Nucleotide sequence and characterisation of the broad-host-range lactococcal plasmid pWV01. Plasmid **26**: 55-66.
Leenhouts, K.J. and G. Venema. (1993). Molecular cloning and expression in *Lactococcus*. In: Proceedings part II Sixth Forum for Applied Biotechnology (Gent, België). pp. 2031-2043.
Leenhouts, K.J., J. Kok and G. Venema. (1991). Replacement recombination in *Lactococcus lactis*. J. Bacteriol. **173**: 4794-4798.
Leenhouts, K.J., J. Gietema, J. Kok and G. Venema. (1991). Chromosomal stabilization of the proteinase genes in *Lactococcis lactis*. Appl. Environ. Microbiol. **57**: 2568-2575.
Leenhouts, K.J. and G. Venema. (1993). Lactococcal plasmid vectors. In: Plasmids. A practical approach. (K.G. Hardy, ed.). IRL Press pp. 65-94.
Leenhouts, K.J. and G. Venema. (1992). General and special purpose lactococcal plasmid vectors. In: Plasmids, a practical approach. (K.G. Hardy, Ed.), Oxford University Press (in press).

Leer, R.J. and J.M.B.M. van der Vossen, M. van Giezen, H.M. van Noort and P.H. Pouwels. (1994). Acidocin B, a non-lantibiotic bacteriocin from *Lactobacillus acidophilus* appears to be processed by a signal sequence dependent peptidase. Microbiol. (submitted).
Lucey, M. (1992). Analysis of the genetic, biochemical and conjugative properties of the phage resistance plasmid pCI528 from *Lactococcus lactis* subsp. *cremoris* UC503. Ph.D. Thesis, National University of Ireland, Cork, IE.
Lucey, M., C. Daly and G.F. Fitzgerald. (1993). Identification and sequence analysis of the replication region of the phage resistance plasmid pCI528 from *Lactococcus lactis* subsp. *cremoris* UC503. FEMS Microbiol Lett. **110**: 249-256.
Lucey, M., C. Daly and G.F. Fitzgerald. (1993). Relationship between *Lactococcus lactis* subsp. *lactis* 952, an industrial starter culture and strains 712, ML3 and C2. J. Appl. Bacteriol. **75**: 326-336.
Lucey, M., C. Daly and G.F. Fitzgerald. (1993). Analysis of the replication and mobilization properties of the phage resistance plasmid pCI528 from *Lactococcus lactis* subsp. *cremoris* UC503. FEMS Microbiol. Rev. **12**: Abst F15.
Lucey, M., C. Daly and G.F. Fitzgerald. (1992). Cell surface characteristics of *Lactococcus lactis* harbouring pCI528, a 46 kb plasmid encoding inhibition of bacteriophage adsorption. J. Gen. Microbiol. **138**: 2137-2143.
Lucey, M., C. Daly and G.F. Fitzgerald. (1993). Analysis of a region from the bacteriophage resistance plasmid pCI528 involved in its conjugative mobilisation between *Lactococcus* strains. J. Bacteriol. **175**: 6002-6009.
Martínez, B., J. E. Suárez and A. Rodríguez. (1994). Nisin Z is the predominant antimicrobial activity of lactococci isolated from an artisanal cheese. J. Dairy Sci. (in preparation).
Mayo, B., P. Arca, B. González and J. E. Suárez. (1994). Cloning and expression of the plasmid-encoded β-galactosidase gene from a *Lactobacillus plantarum* strain of dairy origin. FEMS Microbiol. Lett. (submitted).
McGarry, A., J. Law, A.G. Coffey, C. Daly, P.F. Fox and G.F. Fitzgerald. (1994). Effect of genetically modifying the lactococcal proteolytic system on ripening and flavour development in Cheddar cheese. Appl. Environ. Microbiol. (submitted).
McGarry, A. (1993). The effect of genetically modifying the lactococcal proteolytic system on ripening and flavour development in Cheddar cheese. M.Sc. Thesis, National University of Ireland, Cork, IE.
Mendes-Faia, A., P. Rui Ramos and P. Petta. (1994). Chemical and Sensorial characterisation of wines inoculated with lactic acid bacteria. (in preparation).
Mendes-Faia, A., M. Serpa Pimentel and M. Jose Carneiro. (1994). Tolerance to ethanol of lactic acid bacteria isolated from wine. Symposium of Microbial Fermentations: beverages, foods and feeds, Summer Conference, Herriot-Watt University, Edinburgh, 13-15 July 1994.
Meyer-Barton, E., J.R. Klein, M. Imam, and R. Plapp. (1993). Purification, characterisation, and sequence analysis of X-prolyl-dipeptidyl-aminopeptidase (PepX) from *Lactobacillus delbrückii* ssp. *lactis* DSM 7290. Appl. Microbiol. Biotech. **40**: 82-89.
Monnet, V., M. Nardi, A. Chopin, M.C. Chopin and J.C. Gripon. (1994). Biochemical and genetic characterisation of *pepF*, an oligoendopeptidase from *Lactococcus lactis*. J. Mol. Biol.
Montano, A., M. Bobillo and V.M. Marshall. (1993). Effect of sodium chloride on the metabolism of two strains of *Lactobacillus plantarum* isolated from fermenting green olives. Lett. Appl. Microbiol **16**: 315-318.
Movahedi, S. and F. Mulholland. (1994). Purification and characterisation of a dipeptidase from *Lactococcus lactis* subsp. *lactis* NCDO 712. (in preparation).
Mulholland, F., S. Movahedi, G. R. Hague and T. Kasumi. (1993). Monitoring tripeptidase activity using capillary electrophoresis: A comparison with the ninhydrin assay. J. Chromatography **636**: 63-68.
Nyengaard, N., F.K. Vogensen and J. Josephsen. (1993). LlaA1 and LlaB1, two type-II restriction endonucleases from *Lactococcus lactis* subsp. *cremoris* W9 and W56 recognizing, respectively, 5'-/GATC-3' and 5'-C/TRYAG-3'. Gene. **136**: 371-372.
Payne J. and M.J. Gasson. (1994). Cloning and characterisation of a *Listeria* bacteriophage lysin gene: its potential as a novel antimicrobial agent. Biotechnology. (submitted).
Parreira, R., R. Valyasevi, S.D, Ehrlich and M.C. Chopin. (1994). abi-416 (abiB) encoded abortive infection mechanism induces degradation of lactococcal phage RNAs. (in preparation).

Payne, J. and M.J. Gasson. (1992). Cloning and characterisation of a lysin gene from a *Listeria* bacteriophage. In: NATO ASI Series Vol. H 65 Bacteriocins, microcins and lantibiotics (R. James, C. Lazdunski, F. Pattus Eds.) Springer-Verlag Berlin Heidelberg.
Platteeuw, G., and W.M. de Vos. (1992). Location, characterisation and expression of the lysin encoding gene, *lytA*, of *Lactococcus lactis* bacteriophage ΦUS3. Gene **118**: 115 120.
Poolman, B., E.R.S., Kunji, A. Hagting, V. Juillard and W.M. Konings. (1994). The proteolytic pathway of *Lactococcus lactis*. J. Appl. Bacteriol. (in press).
Pot, B., P. Vandamme and K. Kersters. (1994). Analysis of electrophoretic whole organism protein fingerprints. In: Chemical methods in prokaryotic systematics. (M. Goodfellow and A.G. O'Donnell Eds.) J. Wiley and Sons Ltd. Chichester, U.K. pp. 493 521.
Pot, B. and D. Janssens. (1993). The potential role of a culture collection for identification and maintenance of lactic acid bacteria. In: The lactic acid bacteria. (E.L. Food, H.G. Griffin, R. Molby and C.G. Heden Eds). Proceedings of the first lactic computer conference. Horizon Scientific Press, Norfolk. (Chapter 15, pp. 81-87).
Pot, B., L.A. Devriese, J. Hommez, K. Vandemeulebroecke, K. Kersters and F. Haesebrouck. (1994). Characterisation of *Vagococcus fluvialis* strains isolated from lesions in pigs and cattle and from tonsils of different domestic animals. J. Appl. Bacteriol. (in press).
Pouwels, P.H. and R.J. Leer. (1993). Genetics of lactobacilli: plasmids and gene expression. Antoine van Leeuwenhoek. **64**: 85-107.
Rauch, P.J.G. and W.M. de Vos. (1992). Transcriptional regulation of the Tn5276-located *Lactococcus lactis* sucrose operon. Gene **121**: 55-61.
Renault, P., C. Delorme, J. J. Godon, and S. D. Ehrlich. (1993). Genetique et evaluation du potentiel genetique des souches. Le Lait. **73**: 153-161.
Renault, P., J.J. Godon, N. Goupil, C. Delorme and S.D. Ehrlich. (1994). Metabolic operons in Lactococci. In: Genetics of Streptococci, Enterococci and Lactococci. (J.J. Ferretti Eds.) (in press).
Renault, P., J.F. Nogrett, N. Galleron, J.J. Godon and S.D. Ehrlich. (1994). Specificity of insertion of transposon Tn545 in *Lactococcus lactis* subsp. *lactis*. In: Genetics of Streptococci, Enterococci and Lactococci, (J.J. Ferretti, Eds.) (in press).
Reniero R., P.S. Cocconcelli, V. Bottazzi and L. Morelli. (1992). High frequency of conjugation in *Lactobacillus* mediated by an aggregation-promoting factor. J. Gen. Microbiol. **138**: 763-768.
Reniero R., A. Piva, L. Morelli, V. Bottazzi, and P.S. Cocconcelli. (1993). Purification of *Lactobacillus* secreted proteins. Biotech. Techniques **7**: 401-406.
Rodriguez, J.M., L.M. Cintas, P. Casaus, A. Suarez and P.E. Hernandez. (1994). Detection of lactocin S-producing lactobacilli by PCR. J. Appl. Bacteriol. (submitted).
Rodriguez, J.J., O.J. Sobrino, W.L. Moreira, L.M. Cintas, M.F. Fernandez, B. Sanz and P.E. Hernandez. (1991). Selected characteristics of a bacteriocin from *Lactobacillus sake* 449. Proceedings of the 37th International Congress of Meat Science and Technology, Kulmbach, Germany. pp. 633-636.
Salema, M., B. Poolman, J.S. Lolkema, M.C. Loureiro-Dias and W.N. Konings. (1994). Uniport of monoanionic L-malate in membrane vesicles from *Leuconostoc oenos*. Eurß. J. Biochem. (submitted).
Salema, M., J.S. Lolkema, M.V. San-Romao and M.C. Loureiro Dias. (1994). The proton motive force generated in *Leuconostoc oenos* by L-malate fermentation. J. Bacteriol. (in preparation).
Sarra P.G., L. Morelli and V. Bottazzi. (1992). The lactic microflora of fowl. In: Lactic acid bacteria in nutrition and health. (B. Wood, Ed.) Elsevier Science Publishers Ltd. Barking. U.K. pp. 3-19.
Schouler, C., S. D. Ehrlich and M.C., Chopin. (1994). Sequence and organisation of the lactococcal prolate-headed bIL67 phage genome. (submitted).
Schleifer, K.H., W. Ludwig, R.I. Amann, C. Hertel, M.Ehrmann, G. Kohler and A. Krause. (1992). Phylogenetic relationships of lactic acid bacteria and their identification with nucleic acid probes. In: Actes du Colloque. Lactic 91. Les bacteries lactiques. G. Novel and J.F. Querler Eds.). Centre de Publications de l'Universite de Caen. pp. 23-32.
Seegers, J.F.M.L. (1994). Replication and stability of lactococcal plasmids. PhD Thesis. University of Groningen, NL.
Seegers, J.F.M.L., S. Bron, S., C.M. Franke, G. Venema, and R. Kiewiet. (1994). The majority of lactococcal plasmids carry a highly related replicon. Microbiol. (in press).

Serpa Pimentel, M., M. da Silva, I. Cortes and A. Mendes-Faia. (1994). Growth and metabolism of sugar and acids by *Leuconostoc oenos* under different conditions of temperature and pH. J. Appl. Microbiol. **76**: 42-48.
Serpa Pimentel, M. (1992). Metabolism of citrate and malate by *Leuconostoc oenos*. Project presented as partial fulfilment of the degree in Enology. UTAD, 53 pp.
Shearman, C. A., K. Jury and M.J. Gasson. (1992). Autolytic *Lactococcus lactis* expressing a lactococcal bacteriophage lysin gene. Biotech. **10**: 196-199.
Shearman, C. A., M.-C. Fabri, D. Glazer and M.J. Gasson. (1994). *Lactococcus lactis* subsp. *lactis* aggregation plasmid pAG712 shows homology with *cluA*, a sex factor aggregation gene linked with conjugation. Plasmid (submitted).
Shearman, C. A., K. Jury and M.J. Gasson. (1994). Controlled expression of a *Lactococcus lactis* bacteriophage lysin encoded by two overlapping genes. Appl. Environ. Microbiol. (in press).
Siezen, R.J., W.M. de Vos, J.A.M. Leunissen and B. W. Dijkstra. (1991). Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. Protein Eng. **4**: 719-737.
Siezen, R.J., P.G. Bruinenberg, P. Vos, I.J. van Alen-Boerrigter, M. Nijhuis, A.C. Alting, F.A. Exterkate and W.M. de Vos. (1993). Engineering of the substrate binding region of the subtilisin-like, cell-envelope proteinase of *Lactococcus lactis*. Protein Eng. **6**: 927-939.
Simons, G., M. Nijhuis and W.M. de Vos. (1993). Integration and gene replacement in the *Lactococcus lactis lac* operon: Induction of a cryptic phospho-β-glucosidase in LacG-deficient strains. J. Bacteriol. **175**: 5168-5175.
Sobrino, O.J., J.M. Rodriguez, W.L. Moreira, M.F. Ferandez, B. Sanz and P.E. Hernandez. (1991). Antimicrobial activity of *Lactobacillus sake* isolated from dry fermented sausages. Int. J. Food Microbiol. **13**: 1-10.
Sobrino, O.J., J.M. Rodriguez, W.L. Moreira, L.M. Cintas, M.F. Fernandez, B. Sanz and P.E. Hernandez. (1992). Sakacin M a bacteriocin-like substance from *Lactobacillus sake* 148. Int. J. Food Microbiol. **16**: 215-225.
Stefanitsi, D., L. Psaridi, S. Kyriakidis and G. Sakellaris. (1994). *L. bulgaricus* proteinase induction as a function of protein availability in the sustrate. (submitted).
Stolz, P., G. Bocker, R. F. Vogel and W. P. Hammes. (1993). Utilisation of maltose and glucose by lactobacilli isolated from sourdough. FEMS Microbiol. Lett. **109**: 237-242.
Stolz, P., R.F. Vogel and W. P. Hammes. (1994). Utilization of electron acceptors by lactobacilli isolated from sourdough: II. *Lactobacillus pontis, L. reuteri, L. amylovorus,* and *L. fermentum.* (submitted).
Stolz, P., G. Bocker, W.P. Hammes and R. F. Vogel. (1994). Utilization of electron acceptors by lactobacilli isolated from sourdough: I. *Lactobacillus sanfrancisco* (submitted).
Stolz, P., R.F. Vogel and W. P. Hammes. (1994). Maltose phosphorylase is the key enzyme for the prevalence of heterofermentative lactobacilli in traditionally treated sourdoughs. (in preparation).
Stroman, P., S. Movahedi and F. Mulholland. (1994). Cloning and sequencing of the dipeptidase gene, *pep R*, from *Lactococcus lactis*. (in preparation).
Strøman, P. (1992). Sequence of a gene (*lap*) encoding a 95.3-kDa aminopeptidase from *Lactococcus lactis* ssp. *cremoris* Wg2. Gene **113**: 107-112.
Swindell, S. (1993). The influence of conjugation and transposition on *Lactococcus lactis* genome organisation. PhD. Thesis, University of East Anglia, UK.
Swindell, S., H.G. Griffin and M.J. Gasson. (1994). Cloning, sequencing and comparison of two lactococcal *ldh* genes. Microbiology **140**: 1301-1305.
Tan, P.S.T., B. Poolman, and W.N. Konings. (1993). Proteolytic enzymes of *Lactococcus lactis*. J. Dairy Res. **60**: 269-286.
Tichaczek, P. S., R. F. Vogel and W. P. Hammes. (1993). Cloning and sequencing of *curA* encoding curvacin A, the bacteriocin produced by *Lactobacillus curvatus* LTH 1174. Arch. Microbiol. **160**: 279-283.
Tichaczek, P. S., R. F. Vogel and W. P. Hammes. (1994). Cloning and sequencing of *sakP* encoding sakacin P, the bacteriocin produced by *Lactobacillus sake* LTH 673. Microbiology. **140**: 361-367.
Twomey, D.P., R. Davis, C. Daly and G.F. Fitzgerald. (1993). Sequence of the gene encoding a second *Scr*FI m5C methyltransferase of *Lactococcus lactis*. Gene **136**: 205-209.

Twomey, D.P., R. Davis, C. Daly and G.F. Fitzgerald. (1993). Characterisation of pCI932m harbouring a second m5C methylase associated with the *Scr*FI restriction-modification system. FEMS Microbiol. Rev. **12**:Abst F13.
Van Alen-Boerrigter, I.J., R. Baankreis and W.M. de Vos. (1991). Characterisation and over-expression of the *Lactococcus lactis pepN* gene and localization of its product, aminopeptidase N. Appl. Environ. Microbiol. **57**: 2555-2561.
Van Belkum, M.J. (1991). Lactococcal bacteriocins: Genetics and mode of action. PhD Thesis. University of Groningen, NL.
Van Belkum, M.J., B.J. Haijema, R.E. Jeeninga, J. Kok and G. Venema. (1991). Organization and nucleotide sequence of two lactococcal bacteriocin operons. Appl. Environ. Microbiol. **57**: 492-498.
Van Belkum, M., J. Kok, and G. Venema. (1992). Cloning, sequencing, and expression in *Escherichia coli* of *lcnB*, a third bacteriocin determinant from the lactococcal bacteriocin plasmid p9B4-6. Appl. Environ. Microbiol. **58**: 572-577.
Van de Guchte, M. (1991). Heterologous gene expression in *Lactococcus lactis*. PhD Thesis. University of Groningen, NL.
Van de Guchte, M., J. Kok and G. Venema. (1991). Distance-dependent translational coupling and interference in *Lactococcus lactis*. Mol. Gen. Genet. **227**: 65-71.
Van de Guchte, M., T. van der Lende, J. Kok and G. Venema. (1991). A possible contribution of mRNA secondary structure to translation initiation efficiency in *Lactococcus lactis*. FEMS Microbiol. Lett. **81**: 201-208.
Van de Guchte, M., J. Kok, and G. Venema. (1992). Gene expression in *Lactococcus lactis*. FEMS Microbiol. Rev. **88**: 73-92.
Van de Guchte, M., F.J. van der Wal, J. Kok, and G. Venema. (1992). Lysozyme expression in *Lactococcus lactis*. Appl. Microbiol. Biotechnol. **37**: 216-224.
Van de Guchte, M., E.K. Arendt, J. Horgan, A. Healy, C. Daly and G.F. Fitzgerald. (1993). Molecular characterisation of the temperate lactococcal bacteriophage Tuc2009. FEMS Microbiol. Rev. **12**:Abst F11.
Van de Guchte, M., C. Daly, G.F. Fitzgerald and E.K. Arendt. (1994). Identification of the putative repressor-encoding gene *cI* of the temperate lactococcal bacteriophage Tuc2009. Gene (in press).
Van de Guchte, M., C. Daly, G.F. Fitzgerald and E.K. Arendt. (1994). Identification of *int* and *attP* on the genome of lactococcal bacteriophage Tuc2009 and their use for site-specific plasmid integration in the chromosome of Tuc2009-resistant *Lactococcus lactis* MG1363. Appl. Environ. Microbiol. 60 (in press).
Van der Meer, J.R., J. Polman, M.M. Beerthuyzen, R.J. Siezen, O.P. Kuipers and W.M. de Vos. (1993). Characterisation of the *Lactococcus lactis* nisin A operon genes *nisP* encoding a subtilisin-like serine protease involved in processing and *nisR* encoding a regulatory protein involved in nisin biosynthesis. J. Bacteriol. **175**: 2578-2588.
Van der Vossen, J.M.B.M., J. Kodde, A.J. Haandrikman, G. Venema, and J. Kok. (1992). Characterisation of transcription-initiation and termination signals of the proteinase genes of *Lactococcus lactis* Wg2 and enhancement of proteolysis in *L. lactis*. Appl. Environ. Microbiol. **58**: 3142-3149.
Van der Vossen, J.M.B.M., M.H.M. van Herwijen, R.J. Leer, B. ten Brink, P.H. Pouwels and J.H.J. Huis in't Veld. (1994). Production of acidocin B, a bacteriocin of *Lactobacillus acidophilus* M46 is a plasmid encoded trait: Plasmid curing, genetic marking by in vivo plasmid integration and gene transfer. FEMS Microbiol. Lett. **116**: 333-340.
Van Rooijen R.J., S. van Schalkwijk and W.M. de Vos. (1991). Molecular cloning, characterisation, and nucleotide sequence of the tagatose 6-Phosphate pathway gene cluster of the lactose operon of *Lactococcus lactis*. J. Biol. Chem. **266**: 7176-7181.
Van Rooijen R.J. and W.M. de Vos. (1993). Lysines 72, 80, 213, and aspartic acid 210 of the *Lactococcus lactis* LacR repressor are involved in the response to the inducer tagatose-6-phosphate leading to induction of *lac* operon expression. Protein Eng. **6**: 210-206.
Van Rooijen, R.J. (1993). Characterisation of the *Lactococcus lactis* lactose genes and regulation of their expression. PhD Thesis. Agricultural University of Wageningen, NL.
Vaughan, E., C. Daly and G.F. Fitzgerald. (1992). Identification and characterisation of helveticin V-1829, a bacteriocin produced by *Lactobacillus helveticus* 1829. J. Appl. Bacteriol. **73**: 299-308.

Vaughan, E.E. (1991). Characterisation of a bacteriocin produced by *Lactobacillus helveticus* and cloning of chromosomal genes by heterologous complementation. PhD Thesis. National University of Ireland, Cork, IE.
Veiga da Cunha, M., H. Santos and E.V. Schaftingen. (1993). Pathway and regulation of erythritol formation in *Leuconostoc oenos*. J. Bacteriol. **175**: 3941-3948.
Veiga da Cunha, M., P. Firme, M.V. San-Romao and H. Santos. (1992). Application of 13C-NMR to elucidate the unexpected biosynthesis of erythritol by *Leuconostoc oenos*. Appl. Environ. Microbiol. **58**: 2271-2279.
Venema, G. (1993). Molecular biology and genetic modification of lactococci. J. Dairy Sci. **76**: 2133-2144.
Vogel, R. F., G. Bocker, P. Stolz and W. P. Hammes. (1993). Ecology and metabolism of microbes in sourdough. Book of Abstracts, Microbes in Engineered Ecosystems, April 1993, Wageningen, NL.
Vogel, R. F., B. S. Pohle, P. S. Tichaczek and W. P. Hammes. (1993). The competitive advantage of *Lactobacillus curvatus* LTH 1174 in sausage fermentations is caused by formation of curvacin A. System. Appl. Microbiol. **16**: 457-462.
Vongerichten K.F., J.R. Klein, H. Matern and R. Plapp. (1994). Cloning and nucleotide sequence analysis of *pepV*, a carnosinase gene from *Lactobacillus delbrückii* subsp. *lactis* DSM7290 and partial characterisation of the enzyme. Microbiol. (in press).
Warner P.J. (1993). Microbes in Agriculture. Science Progress. **76**: 1-13.
Warner, P.J. (1993). Genetic Engineering of Microorganisms for Food Production. Conference Proceedings, Food Ingredients Europe, 1993, pp. 232-237. Expoconsult, Maarssen, Netherlands.
Wohlrab, Y. and W. Bockelmann (1992). Purification and characterisation of a dipeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*. Int. Dairy J. **2**: 345-361.
Wohlrab, Y. and W. Bockelmann (1993). Purification and characterisation of a second aminopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*. Int. Dairy J. **3**: 685-701.
Wohlrab, Y. and W. Bockelmann (1994). Purification and characterisation of a new aminopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*. Int. Dairy J. **4**: 409-427.
Wolinowska, R., P. Ceglowski, J. Kok and G. Venema. (1991). Isolation, sequence and expression in *Escherichia coli, Bacillus subtilis* and *Lactococcus lactis* of the DNase (streptodornase)-encoding gene from *Streptococcus equisimilis* H46A. Gene **106**: 115-119.

Patents.

Boot, J.H.A., I. Duitz, M.Y. Toonen, K.J. Leenhouts, A.M. Ledeboer (priority date 14-5-1993). Modification of peptide extracts by a transformed lactic acid bacterium (to be filed).
Chapot-Chartier, M.P., M. Nardi, M.C. Chopin, A. Chopin, J.C. Gripon and V. Monnet (August 13, 1992). Clonage du gene de structure d'une aminopeptidase de *Lactococcus lactis*, production et utilisation de ladite aminopeptidase. French Patent No 2 694 766.
Cluzel, P.J. and M.C. Chopin. (1991). DNA fragments coding for a bacteriophage-resistance mechanism. International patent No WO92/05260.
Cluzel, P. J., and M.-C. Chopin. (1992). DNA fragments coding for a bacteriophage-resistance mechanism. International Patent No W092/05260.
Fitzgerald, G.F. and C. Daly. (1993). Plasmids and uses thereof. European Patent No 87304557.9.
Godon, J.J., P. Renault, S.D. Ehrlich. (1992). Acide nucléique codant pour une α-acetolactate synthase et ses application. French Patent No 11470, extension PCT/GB 93/02012.
Godon, J.J., P. Renault, S.D. Ehrlich. (1992). Acide nucleique codant pour une α-acetolactate decarboxylase et ses applications. French Patent No 11471, extension PCT/FR 93/00909.
Godon, J.J., P. Renault, and S. D. Ehrlich. (September 1992). Acide nucleique codant pour une acetolactate synthase de *Lactococcus lactis* subsp. *lactis* et ses applications. French Patent No 11 470, Extension PCT/GB 93/02012.
Godon, J.J., P. Renault, and S. D. Ehrlich. (September 1992). Acide nucleique codant pour une acetolactate decarboxylase de *Lactococcus lactis* subsp. *lactis* et ses applications. French Patent No 11 471, Extension PCT/FR 93/00909.
Godon, J.J., P. Renault, D. Ehrlich and M.J. Gasson. (1993). Nucleic acid coding for an α-acetolactate synthase and its applications. Patent Application PCT/GB93/02012.
Gruss, A. and E. Maguin. (1992). Plasmides thermosensibles pour la modification genetique des bacteries Gram-positives. Brevet francais 92 03034.

Gruss, A. (1993). Procede de decontamination d'un milieu de fermentation et produit permettant sa mise en oeuvre. Demande de Brevet francais No 93 01776 deposee le 17-02-93.
Klein, J.R. and R. Plapp. Vectors containing nucleotide sequences of a plasmid isolated from *Lactobacillus curvatus*, proteins coded by the genome of these vectors, and the use of such vectors in processes for the production of polypeptides. June 1993. European patent application No. 93 401 585.0.
Klein, J.R., B. Henrich and R. Plapp. Lys-Aminopeptidase PepN from *Lactobacillus delbrückii* ssp. *lactis*, nucleic acids coding for it, and its use in fermentation processes. July 1993. European patent application No. 93 13586.1.
Leenhouts, K.J., J.D. Marugg, C.T. Verrips. (1992). A food grade vector. European Patent application No 0 487 159.
Meyer, B. Henrich, J.R. Klein and R. Plapp. X-Prolyl-Dipeptidyl-Aminopeptidase from *Lactobacillus delbrückii* ssp. *lactis*, nucleic acids coding for the same and its use in fermented food-stuff preparation processes. Dec 1992. European patent application No. 92 403 582.7.
Morelli, L. *et al.* (1993). Pharmaceutical composition comprising selected *Lactobacillus* strains. United States Patent No 5,176,911.
Seegers, J.F.M.L., R. Kiewiet, G. Venema, and S. Bron. (1993). Production of desired proteins or polypeptides by culturing a transformed lactic acid bacterium. European patent No T.7022 (R) — EP.
Unilever patent applicaiton. *Lactobacillus sake*-like strains, production and use of their exopolysaccharides. Priority date 2-12-1993. To be published in June 1994. PCT Appl. EP/93/03338.

T-PROJECT

'HIGH RESOLUTION AUTOMATED MICROBIAL IDENTIFICATION (HRAMI)'

High resolution automated microbial identification (HRAMI) (BIOT CT-910294)

COORDINATOR:
01. K. TIMMIS, Nat. Research Centre for Biotech., GBF, Braunschweig, D
PARTICIPANTS:
02. A. AKKERMANS, Wageningen Agric. Univ., AUW, Wageningen, NL
03. D. BITTER-SUERMANN, Med. Univ. Hannover, Hannover, MHH, D
04. D. COLLINS, Institute of Food Research, Reading, IFR, GB
05. K. SCHLEIFER, Technical University of Munich, TUM, Munich, D
06. P. GRIMONT, Institut Pasteur, PAS, Paris, F
07. M. HÖFLE, Nat. Research Centre for Biotech., GBF, Braunschweig, D
08. K. KESTERS, Rijksuniversiteit Gent, RUG, Gent, B
09. R. de WACHTER, University of Antwerpen, UIA, Antwerpen, B
10. N. KENNEDY, Computer Applied Technologies, CAP, Dublin, IRL

BACKGROUND INFORMATION

A multidisciplinary research team developed automated polyphasic technologies in the fields of molecular genetics, immunology, analytical chemistry, instrumentation and separation science, and applied these technologies to the rapid, accurate identification of microorganisms.

1) Ribosomal RNA (rRNA) gene sequences,
2) chemical 'biomarkers',
3) macromolecular 'profiles' and
4) stable antigens were used to characterize commercially important and environmentally relevant standard sets of microorganisms.

The results were compared and evaluated with respect to identification utility and analytical facility. Those analytical systems found to be useful were used to expand taxonomic data bases and were evaluated in terms of their applicability to the identification, quantitation, and sensitive detection of microbes from environmental samples.

1) Ribosomal (rRNA) RNA gene sequences
(AUW, GBF, IRF, MHH, PAS, TUM, UIA)

In the process of evolution, some genes have been more conserved in their sequence than others. In particular, 16S and 23S ribosomal RNAs act as reliable molecular chronometers and are currently the most useful molecules for estimating phylogenetic relationships of microorganisms. Different parts of the rRNA molecules show considerable variation in the degree of sequence conservation, thereby enabling the use of the molecule to cover a wide range of genealogical relations. In addition to being the most powerful phylogenetic markers known to date, the enormous potential of these molecules for the design of highly specific oligonucleotide probes for microbial identification and detection has recently become apparent. An extensive data base of 16SA and 23S rRNA sequences is required and homologous regions of the sequences need to be compared for successful and economic probe construction. Although the alignment of conserved regions can be performed by programmes currently available, alignment of the variable sequence stretches, which are of special interest for probe design of high specificity, usually has to be refined by hand. Thus, computer programmes need to be developed for the alignment of variable sequence stretches. In addition,

higher-order structure elements need to be considered in the probe construction, since secondary or tertiary structures may influence access of the probes. Such regions will also have to be analyzed by the computer programme.

2. Biomarker chemotaxonomy (GBF, RUG)

Rapid advancements in the field of natural product chemistry have resulted in technological developments which can be used for chemotaxonomic applications. Direct structural analysis of microbial natural products with taxonomic significance facilitates studies of microbial physiology and microbial ecology. Incorporation of heavy isotope labeled or radiolabeled isotopes into specific natural products can be used to study metabolic routes as well as fluxes in pure and mixed cultures of microbes, in reactors and in natural communities.

Structural components of microbial cells, i.e., membrane liquids, cell wall components and lipopolysaccharides, are good chemotaxonomic candidates. They are known to contain useful chemotaxonomic information and are relatively stable with respect to physiological stress. Total cellular fatty acid profiles are now routinely used to help identify microorganisms (i.e. the Marion Sasser Hewlett Packard system). This approach has proved successful for several reasons. The analysis is simple, rapid, and semi-automated. This has greatly facilitated the compilation of a database large enough to assure a reasonably high probability that the fatty acid profile of an unknown will match an entry in the database.

3. Macromolecular profiles

The unifying principle of macromolecular fingerprints of profiles is the biochemical separation, usually by electrophoresis, of a large molecules (MW > 10000) to obtain information on their molecular weight distribution. Molecular weight profiles of all major macro-molecules containing sequence information, i.e., DNA, RNA, and proteins, have been used for the identification of microorganisms. The three types of macromolecules occur in different numbers for individual microorganisms and allow identification at different taxonomic levels. These basic differences in the macromolecular inventory of a microbial cell result in different strategies for their molecular analysis. In general, macromolecular profiles enable a rapid chemotaxonomic overview on any set of individual strains, on a relative basis, according to similarities and differences in the fingerprints. Finally, identification of the microorganisms depends on the availability of reference fingerprints.

3.1. rDNA fingerprints (PAS)

Due to the unlimited possibilities to combine different restriction endonucleases, and given the proper electrophoretic separation, every microbial strain may be identified by its restriction fragment length polymorphism (RFLP). Whereas many RFLP patterns may be too complex to be analysed, probing of the fingerprints with conservative gene probes, such as 16S and 23S rDNA, produce a reduced number of bands with and species- or strain-specific patterns.

3.3. Low molecular weight RNA (LMW) **profiles** (GBF)

Stable RNA consists of two sets of RNA: ribosomal (rRNA) and transfer RNA (tRNA). Both sets comprise around 4000-5000 base pairs of conservative genetic information. RRNA oligonucleotide catalogues or sequences have been the most fruitful in obtaining a new framework for microbial taxonomy. On the other hand, RNAs have just recently been exploited for the identification and taxonomy of

microorganisms. The total tRNA pool, comprising around 30 major tRNAs per bacterial strain, is well suited for macromolecular fingerprinting due to the limited number of molecules with different molecular weights (70-100 nucleotides). Genus- and species- specific tRNA bands are detected by high-resolution electrophoresis which allows the identification of environmental bacteria. The LMW RNAs can be analysed directly from natural ecosystems for the detection of dominant taxonomic groups in natural microbial communities.

3.3. **Protein profiles** *(RUG)*

One-dimensional (1D) electrophoretic separation of cellular proteins of microorganisms is a sensitive technique, mainly providing information on the similarity of strains within the same species or subspecies. Depending on the protein electrophoretic variation within a given taxon, individual strains can often be recognized by small, but specific and reproducible differences in part of their protein fingerprints. In general the technique is simple and unexpensive. Automation and development of improved software is necessary to handle larger numbers of strains.

4. **Immunochemistry** (GBF, MHH)

Serum antibodies have historically been of central importance in the identification of microorganisms. As a result, immunological procedures are well developed, automated and standard in most diagnostic and quality control laboratories, as well as in many research laboratories. The introduction of improved immunochemical procedures into such laboratories should thus be achieved more readily and rapidly than the newer technologies of molecular biology.

The hybridoma technology is a method to obtain populations of homogenous antibodies in unlimited quantities and often, depending on the selection methods, has extremely high specificity. This eliminates some of the problems such as batch-to-batch variability and frequent cross-reactivity of polyclonal antibodies. The standard method of raising monoclonal antibodies by somatic cell hybridization is very time-consuming, and labour-intensive and a major limiting factor in the production of sufficient numbers of monoclonal antibodies for large-scale identification and classification purposes.

The time consuming and labour intensive nature of the hybridoma technique can, in principle, be overcome by use of the new monovalent antibody technology. This involves generating single-chain monoclonal antibodies in *E. coli* through the cloning of the coding sequences of the variable domains of immunoglobulin heavy chans (V_H). This can be accomplished through amplification by PCR (polymerization chain reaction) of the V_H coding sequence using the mRNA of spleens from immunized mice. The cloned V_H domains expressed in *E. coli* allow screening and propagation of monoclonal antibodies within days, rather than the months required to generate conventional monoclonal antibodies. In addition, the lack of need for tissue culture and associated expensive media make automated and efficient screening of monoclonal cDNA libraries in *E. coli* for single domain antibodies very attractive. It should be emphasized that the new technology still guarantees the advantages of the conventional method, i.e., broad varieties of antibodies with high specificity, in unlimited supplies.

OBJECTIVES AND PRIMARY APPROACHES

The major objectives of the proposed research were:

1) to develop high-resolution automated molecular methods for the rapid, accurate, identification of microorganisms;

2) to assess the utility and to compare the efficiency of the different methods developed;

3) to exploit these methods to expand our microbial taxonomy base.

Additionally, the probes (nucleic acid, antibody, etc.) generated for rapid identification and taxonomic purposes will be used for the development of highly sensitive procedures for the detection of specific organisms in environmental samples either *in situ* or following extraction from complex matrices. A further objective is to use the method to assess in contained model systems the possible impact on target microbial communities of the introduction of genetically modified microorganisms into the environment.

RESULTS AND DISCUSSION

A. GBF

1. 16S-RNA sequencing

Sequencing and data analysis have been completed for multiple strains of the core species of *Pseudomonas* listed in Bergey's Manual of Determinative Bacteriology. The 16S rRNA sequences between different species demonstrated differences of approximately 3-7% and are sufficient to differentiate the *Pseudomonas* core species. Additionally *Pseudomonas* species may be subdivided into three groups: Ia, Ib and Ic. The resulting dendrogram is in agreement with the results of DNA-DNA hybridization studies.

In collaboration with TUM, three regions of the 16S rRNA gene sequence which are variable among the *Pseudomonas* species analyzed were identified. These regions, which provide the best target sites for species-specific hybridization probes, are the most diagnostic areas of the 16S rRNA gene, and sequencing of just these regions may be a much more cost-effective means of identifying *Pseudomonas* than sequencing the entire gene.

The first of a series of hybridization probes for the rRNA Group I *Pseudomonas* was been synthesized by TUM, tested for *in situ* applicability, and sent to GBF for specificity screening against the approximately 200 strains of *Pseudomonas* existing in-house.

Complete or partial 16S rRNA gene sequence data has been obtained for approximately 60 additional strains from 20 validly described species of rRNA Group I *Pseudomonas*. Analysis of the three hypervariable regions determined was also carried out to define whether partial 16S rRNA gene sequence determinations are sufficient for the accurate phylogenetic placement of a species of *Pseudomonas*. Phylogenetic relationships of the core species of *Pseudomonas* have been assessed using approximately 1000 nucleotides from the three regions mentioned. The clustering obtained by comparison of partial sequences is essentially the same as that generated by comprison of complete sequences. Sequence analysis of hypervariable regions of more than 150 additional strains is being carried out in collaboration with HRAMI partners MHH and IFR. 16S rRNA gene sequences have been determined for approximately 15 strains of the rRNA Group III (β sub-

group of the *Proteobacteria*) pseudomonads and their phylogenetic relationships defined.

2. Instrumental analysis

a) Supercritical fluid chromatography of respiratory quinones — The more structural isomers of respiratory quinones that can resolve the higher the chemotaxonomic quality of the data. We have found that we can separate more commonly found ubiquinones and menaquinones on polar colums rather than on the nonpolar columns used initially.

b) Infrared spectrometric analyses — Diffuse reflectance infrared Fourier transformed (DRIFT) analysis of the polar lipids (PL) of over 200 strains of RNA group I, II and IV *Psedomonas* was shown to rapidly and non-destructively produce useful chemotaxonomic data. The DRIFT method consistently differentiated *Pseudomonas* strains at the level of RNA groups (e.i. I, II and IV) but was not able to consistently assign strains to RNA group Ia or Ib.

c) HPLC-LSD analysis of intact PL — Laser light scattering detection (LSD) for HPLC has proved to be a sensitive (1 microgram) system for the qualitative and quantitative analysis of intact microbial polar lipids. Using an automated system, the polar lipids of *Pseudomonas* can readily separated into classes based on their polar head groups and the chromatographic data generated can be used ans chemotaxonomic criteria.

d) GC-FID and GC-MS analysis of FAME — Gas chromatographic (GC) analysis of fatty acid methyl esters (FAMEs), derived from the hydrolysis of whole cells, is currently widely used as a rapid means of microbial identification. However, the taxonomic resolution of the method is limited by the quality of the analysis and the system used to reduce and interpret the multivariate data. We have improved the gas chromatographic system for separating FAMEs and verified FAME structures by mass spectrometry. The ability of the improved GC method to resolve bacterial taxa has been assessed by systematically analysing pseudomonads from disparate taxa (rRNA groups I, II and IV) to those known to be more similar (rRNA groups Ia and Ib). FAME data was shown to consistently resolve the major rRNA groups but did not consistently resolve rRNA group Ia from Ib.

3. Phage vector-systems

Two different phage vector-systems were evaluated for the production of antibody fragments in *E. coli*. The first phage display system uses the immunoglobulin variable regions of both heavy and light chain, designated the single-chain Fv fragment, or scFv. The two variable regions are artificially joined together with a linker peptide and expressed as a single polypeptide chain that includes the phage gp3-protein at the C-terminus. The second phage display system involves the larger Fab fragment expressed on the phage surface in combination with gp3. Since both systems have a signal peptide located at the N-terminus of the polypeptide chains the soluble antibody fragments are transported into the periplasm for packaging into phage particles or purification purposes. Starting from mouse hybridomas as a suitable source of mRNA, scFv fragments were successfully cloned and expressed in *E. coli*. Expression levels were determined and shown by immunoblot analyses using a gp3-specific monoclonal antibody. Efforts to incorporate the single chain antibodies in the phage-particle and to use it as an immunological tool are in progress.

B. AUW

The bead-based, selective, 16S ribosomal RNA capture method that we have developed for the analysis of soil samples was tested, optimized and evaluated further. With an overall extraction efficiency of 46% (±10%), the detection limit of the method was found to be 10^9 molecules of a specific 16S ribosomal RNA target (equivalent to approximately 10^4-10^5 cells). Because we are aiming at a specific detection of approximately 10^2-10^4 cells per gram of soil, we need to improve the sensitivity of the method. In order to reach this level of sensitivity it is necessary to utilize nucleic acid amplification techniques. We have focused on the application of the reverse transcriptase polymerase chain reaction (RT-PCR) method. Two different reverse transcriptases (Tth and M-MLV) reverse transcriptase (RT) were tested for their performance. Preliminary results have shown that both the specificity and the sensitivity were higher for Tth RT reactions than for M-MLV RT reactions.

In addition to setting up a selective and generic 16S RNA isolation method that is compatible with RT-PCR, we have developed a rapid, nonisotopic assay for the quantitation of specific RNA species isolated by this method. The assay is based on the quantitave competitive PCR (QC-PCR), and utilizes the semi-automated electro-chemiluminescent QPCR System 5000 (Perkin Elmer) for quantitation of PCR products that are obtained by this method. After isolation using the bead-based selective 16S rRNA capture method, the ribosomal RNA is reverse transcribed, and subsequently coamplified along with an internal standard in one reaction in which the same universal 16S rRNA-targeted primers are used. One of these primers is biotinylated in order to allow capture of the PCR products on streptavidin-coated paramagnetic beads. The internal standard consists of a known natural 16S rRNA sequence that is not present in the ecosystem to be studied. Since the internal standard and the specific 16S ribosomal RNA species share a high degree of sequence homology, are approximately of the same length, and contain the same universal primer binding sites, they are coamplified with the same efficiency during the exponential phase of amplification. This finding can be utilized to quantitate the amount of specific 16S rRNA species by extrapolating against a standard curve that is generated for the internal standard. In a model system, we have demonstrated that when different amounts of 16S rRNA from *Escherichia coli* and *Pseudomonas aeruginosa* are mixed coamplified, the proportionality between the two different targets remains the same.

The complete sequences of the 16S rRNA genes from uncultured symbiotic segmented filamentous bacteria (SFB's) that were isolated from the intestine of mice and rat were determined. The sequences from these different hosts were different at 26 from the 1445 positions that were sequenced, resulting in only 1.8% divergence between these SFB's. The 16S rRNA-targeted probes that were developed based on the sequence analysis of SFB's were tested for their specificity on 25 different *Clostridium* strains that were found to be most related to SFB's. No cross reactivity could be detected, indicating that the probes were specific for SFB's. In addition, we have developed specific 16 rRNA-targeted oligonucleotide probes for the detection of *Clavibacter michiganensis* subsp. *sepedonicus*.

C. MHH

1. Screening of bacterial cells

Monoclonal antibodies have been successfully used for routine analysis in drinking water control. Mab against pan-bacterial antigens (EF-Tu), family-specific

enterobacterial common antigen (ECA), genus-specific heat-shock proteins (HSP), species-specific and strain-specific antigens (mostly carbohydrate in nature, e.g. LPS, capsules, biofilm) are combined in dot blots or ELISAs for a first-line screening. In collaboration with TUM and GSF-Neuherberg we evaluated approaches for combining fluorescently-labelled antibodies and rRNA targeted oligonucleotides for the detection of individual bacterial cells. The combination of these methods should allow an extension of the discriminatory power by using antibodies specific for sub-species determinants which may have been difficult to distinguish by rRNA probes. Another potential advantage of the combined use lies in the detection of organisms under conditions where the amount of ribosomes are not sufficient. We successfully used the mag 898 directly conjugated to FITC with specificity for the Enterobacterial Common Antigen for the identification of cells that before has been hybridized with oligonucleotides. Results demonstrated that microscopic detection was severely influenced by the mode of fixation used to combine these methods. With regard to mab, the detection of e.g. dead bacteria, high cell numbers above 10^4-10^5/ml, all rapid growing bacteria, polysaccharide antigens is possible. In addition, taxonomy at the subspecies level and expression of factors of pathogenicity are domains of the immunochemical method.

2. Automated sequencing procedures

In order to enable an automated sequencing procedure, several steps of the strategy had to be refined. Simplified nucleic acid extraction procedures, and the use of solid phase sequencing methods have been employed to facilitate rapid 16S rRNA sequence determination. In collaboration with IFR the solid phase sequencing strategy was adjusted for automated sequencing procedures. In collaboration with GBF and IFR more than 150 strains of the genus *Pseudomonas* were investigated by rRNA gene sequencing.

In addition, we have focused on the genus *Mycobacterium* in order to evaluate the 16S rRNA gene with respect to identification utility, analytical facility, microbial diversity and characterization of uncultured microorganisms. Our results demonstrate:

a) 16 rRNA sequencing probably represents the single most powerful marker for a genotypic-based approach for identification;

b) 16S rRNA sequencing permits identification of microorganisms which cannot be identified by more traditional techniques; and

c) 16S rRNA sequencing will increase our understanding of microbial diversity, as its discriminating speciation capability far exceeds that of standard biochemical identification techniques. In theory, amplification primers targeting highly conserved regions within the 16S rRNA molecule permit amplification of any microorganisms of interest, including 'noncultivables'.

D. IFR

Work on sequencing of the two highly variable regions of the 23S rRNA from the designated *Pseudomonas* core strains from RNA groups I-IV has been completed. Solid phase automated sequencing strategies developed in conjunction with MHH were used to generate the sequence data. The sequences from the two targeted regions are being assessed and used for probe design in collaboration with TUM.

Primers for the manual sequencing of 23S rRNA, based on universal and semiconserved regions of the molecule, are under further development and assessment. In

conjunction with TUM, a universal set of 23S rRNA primers has been designed for the *Eubacteria* for use in automated sequencing strategies. Four primer pairs have been designed to amplify the 23S rRNA as four overlapping fragments. These include a pair to amplify from the 3' end of the 16S rRNA to the 5' end of the 23S rRNA to allow for the accurate sequencing of the first 100 or so bases of the molecule. The other three primer pairs are designed to generate overlapping fragments of between 800 and 1000 base pairs. These primers have been tested on a range of genera including *Rhizobia, Pediococci* and *Actinomycetes* and produce good amplification products. Biotinylated versions of these PCR primers are now being assessed for solid phase sequencing development.

With some areas of the 23S rRNA molecule a number of different sequencing primer sets are needed to encompass the broadest range of genera. A universal set of primers is also being developed to fill in gaps at positions 600, 1300 and at the 3 end of the molecule. A number of bacterial species have been targeted to obtain full 23S rRNA sequences using the above primers and the solid phase sequencing approach.

A number of complete 16S and 23S rRNA sequences are being determined to increase the number available to the scientific community via the database at the UIA. The 16S rRNA sequences of many members of the genus *Clostridium* have recently been determined at Reading so that all members of the genus have now been covered to determine definitive phylogeny of this important group and to facilitate probe design. In collaboration with GBF and MHH, partial 16S rRNA gene sequencing has been performed on more than 150 strains of the *Pseudomonas rRNA* group I in order to determine the stability of hypervariable regions, selected as targets for oligonucleotide probe design.

The use of rRNA-PCR cloning for determining the geanealogy of non-culturables has been successfully demonstrated with the protozoan parasite *Cryptosporidium*.

E. TUM

1. Software development/rRNA probe database

The ARB program package for handling and analyzing rRNA sequence data designed for SUN Open Windows was further developed. Sequence and additional data can be extracted from public data bases (UIA, RDP, EMBL) by a format converter. Aligned rRNA sequences together with higher order structure information, documentation and a phylogenetic tree are stored in a central database. The tool SEQDIS is used to calculate similarity or distance matrices, profiles and masks (based on overall or positional variability). A preliminary version of the PROBE tool searches potential target sites for specific probes within a defined sequence and evaluates them by comparison with the complete or partial databases. A tool for consensus analysis to find group specific structures and potential sequence or alignment errors is under development. Software for rRNA probe database updating, handling and searching is under development. The future goal is to develop a user friendly environment which allows menu-guided handling of data and tools also by non-expert users. Furthermore the package should be made compatible for almost various common computer systems.

2. Sequencing

Complete 16S rRNA gene primary structures of *Paracoccus denitrificans, 'Magnetobacterium bavaricum'*, the strains of *Zoogloea ramigera, Sarcobium lyticum, Pelobacter acetylenicus* and *Pelobacter propionicus* were determined by direct

sequencing of *in vitro*-amplified CNA. The phylogenetic position of *Dictyoglomus thermophilum* was determined based on the 16S rRNA gene sequence. For seven strains of the genus *Virbio* 16 rRNA and partial 23S rRNA sequences have been determined.

3. Development and testing of rRNA-targeted oligonucleoride probes

a) Group-specific oligonucleotide probes
The specificities of oligonucleotide probes designed for Gram-positive bacteria with a high DNA G+C-content (HGCs) and the *Cytophaga/Flavobacteria*-group have been tested by dot-blot hybridization against nucleic acids isolated from LMG strains and used for the *in situ* detection of defined cell populations in activated sludge, biofilms and sediment. Probes for the two major branches of *Archaea* — the *Crenarchaeota* and *Euryarchaeota* have been developed.

b) Species-specific oligonucleotide probes
The 16S rRNA probe target regions for two *Pseudomonas* core strains (*P. stutzeri, P. aeruginosa*) were designed and evaluated. Currently, partial 23S rRNA gene sequences of selected *Pseudomonas* core strains obtained from IFR are used to design an additional set of 23S rRNA-targeted oligonucleotide probes.

4. Combination of fluorescent probes and immunofluorescence

Initial experiments testing the feasibility of combined whole cell probing with antibodies and oligonucleotides have been pursued. A fluorescein-labeled monoclonal antibody directed against the enterobacterial common antigen (ECA) was successfully used for the identification of cells that had been hybridized with oligonucleotides before. Preliminary experiments demonstrated that this combined antibody/oligonucleotide probing is compatible with flow cytometric analysis.

5. Improvement of whole cell hybridization techniques

a) Oligonucleotides
For higher sensitivity we are currently evaluating the use of enzyme-labeled probes for *in situ* hybridization in aquatic environments.

b) Polynucleotides
16S rRNA-targeted transcript probes confer up to 26 times more fluorescence to cells than monolabeled oligonucleotide probes. These probes have been applied to the specific detection of populations in activated sludge.

6. Image analysis

The focus was on rapid means to quantify the fluorescence emitted from individual cells. Some promising progress was achieved. Further work must concentrate on the transfer of user-friendly versions into the hands of applicants.

7. Applications

The *in situ* hybridization technique is now routinely used for the detection of individual cells in activated sludge and biofilms. A combination of direct sequence retrieval and *in situ* probing with specific oligonucleotides facilitated the phylogenetic analysis of hitherto uncultured endosymbionts of protozoa, e.g. of *Caedibacter caryophila*, an obligate endonuclear bacterial symbiont of *Paramecium caudatum*, and of *Polynucleobacter necessarius*, an cytoplasmatic symbiont of *Euplotes aediculatus*.

F. PAS

Automatic DNA extraction and purification: by AutoGen 540 is routinely used to extract and purify 8 DNA samples in 75 minutes. The machine, which is well accepted by the laboratory personnel, produces about 5000 DNA samples per year. Each DNA sample can be used for about 10 restriction experiments.

1. Determination of rRNA gene restriction patterns

Technical improvements yielding well resolved banding patterns included low voltage electrophoresis, vacuum transfer, use of a hybridization incubator, and chemically labeled probe. Probes giving good results were acetylaminofluorene-labeled rRNA or diogoxigenin-labeled oligonucleotide mixtures. Patterns have been determined for several groups of bacteria including the 'core strains' distributed by Partner 8 (RUG). Of 19 restriction endonucleases tested on a few strains, only *Hinc*II, *Sma*I, *Hind*III, and *Sac*I were found to be useful. All *Pseudomonadaceae* strains distributed as a 'core' collection were studied using *Hinc*II and *Sma*I. Most species gave a species-specific pattern with minor variations enabling typing. Exceptions were *P. fluorescens, P. putida, P. chlororaphis* which were extremely heterogeneous.

2. Correlation with phenotypic features

The strains were also studied by carbon source utilization tests using Biotype-100 strips (BioMérieux). A numerical taxonomy of nutritional data (software used: Recognizer Taxolab Institut Pasteur) yielded phenons which could generally match rRNA gene restriction patterns. *P. fluorescens, P. putida, P. chlororaphis* were again found to be heterogeneous. Some species could not be separated from other species by nutritional tests: *P. coronofaciens* from *P. caricapapayae, P. fuscovaginae* from *P. asplenii, P. oleovorans* from *P. resinovorans, P. aureofaciens* from *P. chlororaphis* (type strain), *Acidovorax facilis* from *A. delafieldii.*

3. Image capture

Two image capture systems were tested:

a) Viso-Mic (Genomic, Collonges-sous-Salève, France) included a CCD camera, a 386 PC equipped with a videocapture card, an enhanced VGA monitor and a video copy processor, and
b) One-scanner (Appel Computers), a black-and-white flat-bed scanner. Both were used to capture rRNA gene restriction patterns in the form of a TIFF file.

4. Analysis of captured date

A software package named Taxotron and comprising modules RestrictoScan, RestrictoTyper, Adanson and Dendrograf written by P.A.D. Grimont is now routinely used to read TIFF or PICT files, detect lanes and bands (smoothing available), calculate migration values, interpolate fragment size accurately (Schaffer & Sederoff or Cubic Spline algorithms), provide a schematic representation of fragment migration which can be used in publications, calculate distance matrix and draw rooted or unrooted taxonomic trees. From fragment size (MW) files (which can be edited and merged), databases can be built. Any pattern, described by a set of fragment lengths, can be compared to a database.

5. Ribotype database

Sources of variation in MW determination have been identified. The 1.5% MW error due to manual capture of migration data has been reduced to less than 0.2% (pixel rounding). Between-experiment error has been kept below 3.5% when four standard lanes are used in a 20-lane gel. In some bacterial groups e.g., *Escherichia coli* some endonucleases (e.g. *Mlu*I) yield patterns sharing common bands. These common bands can then be used as internal markers, thus reducing the between-experiment error to less than 2%. Ribotype databases are now available for *Pseudomonadaceae, Flavobacterium-Cytophaga, Escherichia coli, Shigella, Salmonella* ser Typhi, *Citrobacter* spp., *Acinetobacter* spp., *Propionibacterium* spp., *Methylophaga* spp. These databases will be improved and completed regularly.

G. GBF (MPI)

1. Development of high-resolution LMW RNA profiles by conventional electrophoresis

High resolution gels (80 cm, denaturing polyacrylamide) were developed with a resolution of 40-50% more rRNA bands than standard gels. Progress, in terms of purification of rRNA, was made by using purification columns with an ultrafiltration membrane (cut-off 100,000 dalton). A fast screening gel technique was also developed that allows the rapid screening of up to 60 strains on the same high-resolution gel.

2. Comparison of high-resolution LMW RNA profiles of core strains

High-resolution electrophoresis was completed for more than 150 core strains of the HRAMI project. 146 strains of the *Pseudomonas* RNA group I were analyzed for their LMW RNA profiles. The results indicate that all strains, exept four (LMG 2197, LMG 5081, LMG 5824, LMG 5837), are true RNA group I pseudomonads due to their genus specific bands in the tRNA band pattern. From the 142 true *Pseudomonas* core strains, most strains and species could be attributed by their tRNA pattern to one of the two main subgroupings of the genus *Pseudomonas*, i.e. subgroup Ia of the *P. aeruginosa* type or subgroup Ib of the *P. fluorescens* type. A third subgroup could be recognized centred around *P. syringae* with closely related species such as *P. ficuserectae* and *P. caricapapaya*. A last group of the core *Pseudomonas* species analyzed so far were the exceptional species that did not fit in any of the 3 subgroups due to their very different tRNA pattern. These exceptional species were *P. stanierii* and *P. pertucinogena*.

3. Application of LMW RNA profiling to environmental samples and isolates

The overall structure of the bacterioplankton community from the central Baltic was analyzed by high resolution electrophoresis of its LMW RNA fraction. This technique gave a direct overview of the taxonomic diversity of the bacterial community, i.e., the number of taxa and their relative amounts occurring in a sample. For further information about the identity of single taxa partial sequencing of 5S rRNA was used. This enabled a more detailed comparison of specific taxa in samples derived from different depths as well as comparison with reference strains. One isolate was identified as a *Thiobacillus denitrificans*-like species by partial sequencing of the 5S rRNA. With the standard LMW RNA profiling technique a set of 77 nitrous oxide producing isolates from the same water samples of the central Baltic was screened. The bulk of the isolates (almost 80%) could be identified

as *Shewanella putrefaciens*. The rest of the strains (18) belonged to 12 different genotypes often represented by a single isolate.

4. Direct automated 5S rRNA sequencing

Further taxonomic information from mixed cultures or environmental samples can be obtained by analysis of the 5S rRNA bands as a second analytical step after conventional electrophoresis of the LMW RNA. For this purpose, 55 rRNA was transcribed into cDNA, amplified and subjected to cycle sequencing. The resulting complete sequence of the 5S rRNA could be directely fed into any electronic data base and sequence alignment performed to obtain taxonomic and phylogenetic information.

6. Application of capillary electrophoresis to LMW RNA analysis

For the development of an appropriate separation technique of the LMW RNA by capillary electrophoresis two different approaches were explored:
a) capillary gel electrophoresis using polyacrylamide gel-filled capillaries (3% and 5% PAGE); and
b) capillary electrophoresis using non-gel sieving buffer.

The capillary gel electrophoresis showed a good resolution only for small tRNAs (class 1 tRNA to 79nt).

The non-gel sieving capillary electrophoresis showed good separation of the three main LMW RNA groups, i.e. 5S rRNA, class 1 and class 2 tRNA. The three main fractions were separated well either due to different size (1-5 nucleotides) or secondary structure. Good reproducibility in terms of migration times of the main peaks was obtained. Repeated measurements (more than 100) were carried out under the same running conditions. The relative standard deviation of the retention time of the various LMW RNA standards was lower than 0.4% across runs. In comparison with the capillary gel electrophoresis, the non-gel sieving capillary electrophoresis showed good resolution for all LMW RNA fractions, high reproducibility and stability of the analytical system. These features are prerequisites for automation of the method.

6. Establishing of a LMW RNA profile electronic data base

Automated data analysis of the LMW RNA profiles from slab gels was improved by importing the raw data of the gel scanner directly into a data analysis software package. The data analysis package allows an automated search for the most similar reference gel scan in a data base for a new, unidentified strain.

H. RUG

1. Selection and distribution of strains

A total of more than 1100 reference strains (groups: pseudomonads, *Flavobacterium-Cytophaga* complex and enterococci) were distributed to different HRAMI participants (GBF, IFR, MHH, PAS, RUG, TUM). Thirty three new isolates were obtained from GBF, stored in the LMG Culture Collection and sent to PAS and RUG for further characterization.

2. Development and improvement of software

Software was developed for the numerical comparison of BIOLOG data. The software for data capture and alignment of protein profiles was improved.

3. Taxonomic results

Three fingerprinting techniques (SDS-PAGE, FAME and BIOLOG) were applied for the characterization and identification of approximately 275 different strains: 200 reference strains of the pseudomonads, 22 strains of the *Flavobacterium-Cytophaga* complex, 20 core strains of the enterococci and 33 new isolates. The reference strains of the pseudomonads are members of 42 different species of which the majority belongs to the pseudomonads of RNA group 1. Databases were constructed and used for the identification of the new isolates. The level at which the respective techniques yield valuable taxonomic information was evaluated and the accuracy of the commercial databases (FAME and BIOLOG) was determined.

4. Characterization of reference strains

Characteristic fatty acid patterns are obtained for the three microbial groups studied. Within these taxa some major subgroups can be recognized. For the pseudomonads the data indicate that each of the rRNA groups (I to IV) has a characteristic fatty acid profile where the hydroxy acids are of particular taxonomic value. Within RNA group I, groups of species are recognized which may reflect close phylogenetic relationships. Some species, form separate entities (e.g. *P. aeruginosa*) display a considerable heterogeneity in their fatty acid profiles (e.g. *P. stutzeri*). Because of a high similarity among the profiles of several species many reference strains are misidentified when the commercial FAME database is used.

5. Investigation in bacterial phylogeny

a) Investigation of bacterial phylogeny in general

A tree derived from an alignment of 1232 bacterial sequences has been developed. Similar trees, constructed regularly on the growing alignment, showed a reproducible pattern of clusters corresponding to established divisions and subdivisions of the Bacteria. However, the branching pattern leading to these clusters was not stable but varied according to the composition of the species set available at the moment of tree construction. Such a situation is often observed in studies in molecular evolution. As a sequence alignment increases in size, trees based on the alignment shows a reproducible pattern of clusters, e.g. the metazoa, fungi, plants and a number of Protist taxa in the case of the eukaryotic domain. However, the topology of the branching pattern leading to these taxa often changes as a function of the composition of the growing dataset. This has prompted us to develop a method for assessing the stability of branching patterns leading to a set of taxa in phylogenetic trees. It consists in the construction of a large number of trees, each containing a single representative of each taxon. A consensus tree is then derived that shows the frequency of occurrence of each cluster in the tree..

Application of this method to the bacteria led to a consensus tree. Among the 16 clusters observable in a divergence pattern leading to 18 bacterial divisions and subdivisions, only 5 branching points are observed with a frequency of more than 50%. The order of divergence of several bacterial divisions and subdivisions remains indeterminate.

b) Investigation of the phylogeny of* Proteobacteria *β and γ subdivisions

The bacterial phylogenic trees also showed an anomaly in the clusters containing the β and γ subdivisions of the Proteobacteria. The γ cluster seems to behave as a polyphyletic taxon, with a number of genera diverging prior to the divergence among the β subdivision and the remaining genera of the γ subdivision. This phenomenon has also been investigated with the aforementioned program,

however by systematic selection of each member of the Proteobacteria α subdivision as outgroup species.

I. UIA

1. Database on rRNA structure

At present the SSU rRNA database contains 2824 entries and the LSU rRNA database contains 261 entries. The database can be accessed by anonymons ftp and is made available on magnetic media or as hard copy for HRAMI members who do not have access to networks.

2. Software developed for the maintenance and exploitation of the database

a) Alignment editor
The alignment of sequences and the indication of secondary structure patterns on each sequence requires special editing facilities that allow to shift sequence sections individually or in groups, without taking the risk of accidental introduction of errors in the recorded primary structure. A special editor, named DCSE (dedicated comparative sequence editor) has been developed for this purpose. Apart from being essential in the alignment process, the editor is proving very useful for the continuous refinement of the secondary structure pattern encoded in the sequence alignments. Other functions performed by this editor are the search for sequence complementarity and the derivation of conserved sequence sections or consensus sequences which can be used for the design of oligonucleotides serving as hybridization probes or primers for PCR or sequencing.

b) Software for derivation of phylogenetic trees
As mentioned in previous reports, a software package named TREECON has been developed for reconstruction of phylogenetic trees from nucleotide sequence alignments. This package, which runs on IBM-compatible microcomputers, can be obtained by frp or on magnetic disk. A version has been developed that runs on a VaxStation and that can take advantage of the larger memory capacity of a workstation to construct phylogenetic trees from more than 1000 species. These programmes have been used for the investigation of bacterial phylogeny on the basis of the extensive SSU rRNA dataset now available. Phylogenetic trees were constructed by neighbour joining, which is the algorithm that gives the most accurate results while still allowing the handling of large numbers of species.

3. Investigation in bacterial phylogeny

a) Investigation of bacterial phylogeny in general
A trees derived from an alignment of 1232 bacterial sequences has been developed. Similar trees, constructed regularly on the growing alignment, showed a reproducible pattern of clusters corresponding to established divisions and subdivisions of the Bacteria. However, the branching .pattern leading to these clusters was not stable but varied according to the composition of the species set available at the moment of tree construction. Such a situation is often observed in studies in molecular evolution. As a sequence alignment increases in size, trees based on the alignment show a reproducible pattern of clusters, *e.g.* the metazoa, fungi, plants and a number of Protist taxa in the case of the eukaryotic domain. However, the topology of the branching pattern leading to these taxa often changes as a function of the composition of the growing dataset. This has prompted us to develop a method for assessing the stability of branching patterns leading to a set of taxa in phylogenetic trees. It consists in the construction of a large number of

trees, each containing a single representative of each taxon. A consensus tree is then derived that shows the frequency of occurrence of each cluster in the tree. Application of this method to the bacteria led to a consensus tree. Among the 16 clusters observable in a divergence pattern leading to 18 bacterial divisions and subdivisions, only 5 branching points are observed with a frequency of more than 50%. The order of divergence of several bacterial divisions and subdivisions remains indeterminate.

b) Investigation of the phylogeny of **Proteobacteria** *β and subdivisions*

The bacterial phylogenetic trees also showed an anomaly in the clusters containing the β and γ subdivisions of the Proteobacteria. The γ cluster seems to behave as a polyphyletic taxon, with a number of genera diverging prior to the divergence among the β subdivision and the remaining general of the γ subdivision. This phenomenon has also been investigated with the aforementioned programme, however by systematic selection of each member of the Proteobacteria α subdivision as outgroup species.

K. CAP

1. Software prototyping

a) Prototype software tools were developed to address the following main tasks:
   i) Elimination of background noise and artifact from epifluorescence images.
   ii) Elimination of cell clumps and identification of individual cells for counting.
   iii) Measurement of defined cell characteristics and output of results.
   iv) Distinction between cells in the same image, labelled with various ratios of fluorescent labels of two or more different colours.

2. Software evaluation

TUM has reported to CAPTEC that the image analysis with the BIO-MAVIS™ software tools yielded results which compared favourably with standard laboratory methods, where comparable methods were available. Some parameters measurable with image processing could not be compared with existing methods (such as morphometric details). However, determination of cell numbers in bacterial suspensions using image analysis were comparable with FACS results. Determination of the fluorescent intensity of bacterial cell suspensions using image analysis compared well with flow cytometric analysis, although yielding consistently higher intensity measurements, which may reflect the fundamental differences between the two methods used.

3. Implications for CAPTEC

Participation in the HRAMI project gave CAPTEC the opportunity to further develop its BIO-MAVIS™ software library to deal with a new range of applications in epifluorescence microscopy and to address a further array of potential image-processing applications in the life sciences which would not have been possible without participation in an international group having expertise in such a wide spectrum of laboratory techniques. The software provided by CAPTEC to TUM continues to be used in the course of their research using fluorescent-labelled oligonucleotide probes in laboratory cultures of a variety of *Pseudomonas* species, having been found to complement other image-analysis software tools

(OPTIMAS) and to provide their first reliable method of quantifying fluorescent intensity of such cells using image analysis.

The software tools developed by CAPTEC, specifically those used to measure the fluorescent intensity of bacterial cells labelled with fluorescent oligonucleotide probes, represent an extension to existing capability in the analysis of biological images. In addition to the benefits of the quantitation of fluorescence in the identification of microbes in environmental samples (as in the HRAMI project), these new software tools may have potential applications in other fields of biological research where the quantitation of fluorescence in other cell types may be of value. Such application areas include immunology, haematology, genetics and biochemistry, among others.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

- Complete 16S rRNA gene sequences have been determined for all validly-describes *Pseudomonas* spp.
- A new method of supercritical fluid chromatographic analysis of isoprenoid quinones was developed and applied to the analysis of over 100 strains of RNA group I *Pseudomonas*.
- Hypervariable sequence regions that characterize *Pseudomonas* species have been identified and exploited for oligonucleotide probe development.
- DRIFT spectrometry was applied to the analysis of complex bacterial polar lipids and shown to be a potentially rapid method for identification of bacteria to at least the genus level.
- Production of a monoclonal antibody against M13 gp3 which can be used to study gp3-fusion protein expression in phage-display systems.
- Development of a direct and rapid ribosomal RNA extraction method.
- Development of a rapid, sensitive and nonisotopic assay based on the quantitative competitive PCR.
- Characterization of uncultured symbiotic segmented filamentous bacteria (SFB's) inhabiting the small intestine of mice and rat.
- Discovery of a family-specific mab for *Enterobacteriaceae* and genus-specific mabs for *Legionella* and *Pseudomonas*.
- Demonstration of the applicability of a mab against elongation factor TU as a universal antibody against all prokaryotes including *Archaea*.
- Successful design of ELISAs for all 4 antibodies and evaluation to their environmental applicability.
- Establishment of a solic phase sequencing strategy which was adjusted for automated sequencing procedures.
- Sequencing of the two highly variable regions of the 23S rRNA from designated *Pseudomonas* core strains from RNA group I-IV.
- Development of solid phase automated sequencing-strategies.
- Software development for fast handling and analyzing rRNA sequence data for SUN Open Windows.
- Complete 16S rRNA gene sequence of Paracoccus *denitrificans, Magnetobacterium bavarium, Zoogloea ramigera, Sarcobium lyticum, Pelobacteracetylericus, Pelobacter propionii.*
- Design of group-specific oligonucleotide probes (HGCs, Cytophaga/Flavobacteria-group.
- Design of species specific oligonucleotide probes (*Pseudomonas stutzeri, Pseudomonas aeruginosa*).

- The present work is the first application of automated ribotype identification based on molecular sizes of restricted fragments.
- **Commercial availability of the software package.** The Institut Pasteur has made the above-described software package commercially available.
- **Industrial applications.** Ribotyping and the automatic retrieval and identification of ribotypes are now used for the characterization of industrial strains in addition to solving epidemiological problems. The method and software could well be used to compare a strain released in the environment and isolates of the same species that would be isolated from man, animals, plants, or the natural environment.
- Improvement of resolution for tRNAs by 50% and increasing of the rapidity by a factor of 3 for electrophoretic LMW RNA profiles.
- Demonstration of the applicability of LMW RNA profiling to a set of marine samples and bacterial isolates.
- Development of an automated, PCR-driven 5S rRNA sequencing technique.
- Development of a capillary electrophoresis technique for LMW RNA profiling with the potential of automation.
- Establishing of an electronic data base on high resolution LMW RNA profiles.
- Development of software to enhance automation and sensivity or profiling methods.
- Accumulation of extensive chemotaxonomic and phenotypic data bases.
- Comparison of the different techniques for identification of microorganisms.
- Software development for the maintenance and exploitation of the database.
- CAPTEC had had interest in this new software from other research groups using epifluorescence microscopy in Microbiology and from commercial biomedical product developers regarding the use of the software developed in the HRAMI project.

MAJOR COOPERATIVE LINKS

- Sequence analysis of the hypervariable regions of approximately 150 strains carried out in collaboration with HRAMI partners GBF, IFR and MHH.
- Experiences in primer design and phage-display techniques are exchanged between GBF (M. Tesar) and MHH (D. Bitter-Suermann and Köhl).
- Phylogenetic positioning on the basis of 16S rRNA sequence analysis and development of specific rRNA-targeted oligonucleotide probes for detection and identification of uncultured microorganisms is carried out in close collaboration with the group of K.-H. Schleifer and W. Ludwig (TUM) and the group of M.D. Collins (IFR). A reference strain of *Pseudomonas aeruginosa* was provided by E. Moore (GBF), and DNA samples isolated from different *Clostridium* strains were obtained from M.D. Collins for testing the specificity of the developed SFB probes.
- In collaboration with TUM (R. Amann) and GSF-Neuherberg (Dr. Wallner, Dr. Reisker), approaches were evaluated for combining fluorescent-labelled antibodies and rRNA targeted oligonucleotides for the dection of individual bacterial cells.
- MHH (E. Böttger) in collaboration with D. Collins (IFR), solid phase sequencing was adjusted for automated sequencing procedures.
- In collaboration with E. Moore (GBF), D. Collins (IFR) and E. Böttger (MHH) more than 150 strains of the genus *Pseudomonas* were investigated by rRNA sequencing.
- Development of probes for the two major branches *Archaea, Crenarchaeota* and *Euryarchaeota* in collaboration with Prof. Stetter and Dr. Burggraf.

- 16S rRNA probe target regions for *P. stutzeri* and *P. aeruginosa* were evaluated in collaboration with GBF and LMG.
- 16 rRNA and 23S rRNA gene sequences have been deposited with UIA with a central repository.
- Whole cell probing with antibodies and oligonucleotides have been pursued by TUM in collaboration with MHH.
- Image analysis of pure culture has been carried out by TUM in collaboration with CAP.
- Exchange of strains by RUG with GBF, IFR, MHH, PAS, RUG, TUM.
- Exchange and training of a technician from GBF with PAS for DNA-DNA hybridization.
- Collaboration between RUG and PAS in terms of improved software for fingerprint data analysis was very valuable. Comparison of different fingerprinting techniques in respect to taxonomic resolution is done in close collaboration with RUG and PAS.
- The utility and the efficiency of the different profiling methods is evaluated in close collaboration between the partners of GBF, PAS and RUG.
- Organization and access to the 16S rRNA database, by magnetic media or as hard copy, was provided by UIA for HRAMI members.
- Development and optimization of image analyses software coordinated between CAP and TUM.

PUBLICATIONS

Joint publications

Moore, E.R.B., Wittich, R.-M., Fortnagel, P. and Timmis, H.N. (1993). 16S rRNA gene sequence characterization and phylogenetic analysis of a dibenzo-*p*-dioxin-degrading isolate within the new genus *Sphingomonas. Letters in Applied Microbiology* **17**: 115-118.

Shouch, Y, Moore, E.R.B., Ludwig, W., Timmis, K.N. Analysis of genera of the family Comanadaceae by 16S rRNA. (submitted).

Snel, J., Poelman, F.G.J., Koopman, J.P., Blok, H.J., Kengen, H.M.P. and Akkermans, A.D.L. Phylogenetic characterization of *Clostridum* related segmented filamentous bacteria in mice based on 16S ribosomal RNA analysis. (submitted).

Akkermans, A.D.L., Mirza, M.S., Harmsen, H.J.M., Blok, H.J., Herron, P.R., Sessitsch, A. and Akkermans, W.M. Molecular ecology of microbes: a review of promises, pittfalls and real progress. *FEMS Microbiol. Reviews.* (in press).

Mirza, M.S., Rademaker, J.L.W., Danse, J.D. and Akkermans, A.D.L. Specific 16S ribosomal RNA targeted oligonucleotide probe against *Clavibacter michiganinensis* subsp. *sepedonicus. Can. J. Microbiol.* **39**: 1029-1034.

Blok, H.J., Gohlke, A.M. and Akkermans, A.D.L. Quantitation of 16S rRNA-targeted probe signals by the QPCR System 5000. Proceedings of the second annual conference 'New Horizons in Gene Amplification Technologies: New Techniques and Applications'. San Francisco, CA, USA. June 8-10, 1994.

Hutson, R.A., Thompson, D.E., Lawson, P.A., Schocken-Itturino, R.P., Böttger, E.C. and Collins, M.D. (1993). Genetic relationships of *Clostridium botulinum* types A, B and F and other members of the *Clostridium botulinum* complex as revealed by small-subunit rRNA gene sequences. Antonie van Leeuwenhoek **64**: 273-283.

Hutson, R.A., Böttger, E.C., Stubbs, S. and Collins, M.D. Improved manual and automated sequencing strategies for the determination of ribosomal RNA gene sequences. *J. Microbiol. Meth.* (in press).

Vancanneyt, M., Kesters, H., Witt, S. and Fredrickson, H.L. Chemotaxonomy of *Pseudomonas* based on total lipid fatty acids compared to polar lipid acids. (submitted).

Individual publications

Moore, E.R.B., Tindall, B.J., Yakimov, M.M., Fredrickson, H., Vancanneyt, M., Kersters, K. and Timmis, H.N. Characterization of *Pseudomonas* (*sensu strictu*) species by genotypic and chemotaxonomic markers. (submitted).

Bennasar, A., Rossello, R, Lalucat, J., Timmis, K.N., Moore, E.R.B. 16S rRNA gene sequence analysis relative to genomovars of *Pseudomonas stuzeri*: proposal of Pseudomonas baliarica, sp. nov. (submitted).
Asturias, J.A., Moore, E., Yakimov, M.M., Klatte, S. and Timmis, K.N. (1994). Reclassification of the polychlorinated biphenyl-degraders Acinetobacter sp. strain P6 and *Corynebacterium* sp. strain MB1 as *Rhodococcusgloberulus. System. and Appl. Microbiology* **943.** (in press).
Tesar, M., Beckmann, C., Faude, U. and Timmis, K.N. Monoclonal antibody gainst phage M13 gp3: An immunological tool to study gp3 fusion protein gene expression in phage display systems. (submitted).
Steinmetz, I., Rheinheimer, C. and Bitter-Suermann, D. Rapid identification of *Legionellae* by a colony blot assay based on a genus-specific monoclonal antibody. *J. Clin. Microbiol.* **30**: 1016-1010 (1992).
Hübner, I., Steinmetz, I., Obst, U., Giebel, D. and Bitter-Suermann, D. Rapid determination of members of the family *Enterobacteriaceae* in drinking water by an immunological assay using a monoclonal antibody against enterobacterial common antigen. *Appl. Environm. Microbiol.* **58**: 3187-3191 (1992).
Determination of enterobacterial common antigen (ECA) for the detection of lactose fermenting *Enterobacteriaceae*. In: German Standard Methods for the Examination of Water, Waste Water and Sludge; Microbiological Methods (Group K); Prenorm DIN V 38 41 1, part 9, 1993.
Köhl, J., Hartmann, H., Casaretto, M., Köhl, G., Bautsch, W, Kios, A. and Porth, M. Mol. Immunol. **30**: 25 (1993) (Abstract) Characterization of recombinant Fab-fragments, which react with a neoepitope on C3a/C3a desarg.
Steinmetz, I, Rheinheimer, C. and Frosch, M. Application of a genus-specific monoclonal antibody for the detection of *Legionellae* in clinical specimens ans the identification from culture plates. In: Barbaru, J.M., Braman, R.F., Dufour, A.P. (Eds.) Legionella. American Society for Mibrobiology, Washington, D.C., 1993, pp. 22-23.
Steinmetz, I and Bitter-Suermann, D. Monoclonal IgA.-Switch-variants against bacterial respiratory tract pathogens: Molecular forms and transport into murine secretions, (submitted).
Weber, S., Lottspeich, F. and Köhl, J. Detection of an almost panbacterial epitope by a monoclonal antibody to the elongation factor Tu. (submitted).
Kirschner, P. Kiekenbeck, M., Meissner, D., Wolters, J. and Böttger, E.C. (1992). Genetic heterogeneity within *Mycobacterium fortuitum* complex species: genotypic criteria for identification. *J. Clin. Microbiol.* **30**: 2772-2775.
Meier, A., Kirschner, P., Schröder, K.H., Wolters, J. Kroppenstedt, R.M. and Böttger, E.C. (1993). *Mycobacterium intermedium* sp. nov. *Int. J. Syst. Bacteriol.* **43**: 204-209.
Springer, B., Kirschner, G., Rost-Meyer, G., Meier, A., Schröder, K.H., Kroppenstedt, R.M. and Böttger, E.C. (1993). *Mycobacterium interjectum*, a new species isolated from a patient with chronic lymphadenitis. *J. Clin. Microbiol.* **31**: 3083-3089.
Kirschner, P., Springer, B., Bange, F., Wrede, A. Klekenbeck, M., Meier, A., Vogel, U. and Böttger, E.C. (1993). Genotypic identification of mycobacteria by nucleic acid sequence determination. *J. Clin. Microbiol.* **31**: 3083-3089.
Haas, W., Kirschner, P., Butler, R., Plikaytis, B. Salfinger, M., Bremer, H.J. and Böttger, E.C. *Mycobacterium heidelbergense* sp. nov. (submitted).
Meier, A., Persing, D.H., Finken, M. and Böttger, E.C. (1993). Elimination of contaminating DNA within polymerase chain reaction reagents: implications for a general approach for detection of unculturele pathogens. *J. Clin. Microbiol.* **31**: 646-652.
Collins, M.D., Stubbs, S., Hommez, J. and Devrise, L.A. (1993). Molecular taxonomic studies of *Actinomyces-like* bacteria isolated from purulent lesions in pigs. Description of *Actinomyces hyovaginalis* sp. nov. *International Journal of Systematic Bacteriology* **43**: 471-473.
Funke, G., Stubbs, S. and Collins, M.D. (1994). Assignment of human-derived CDC group I coryneform bacteria and CDC group 1-like coryneform bacteria to the genus *Actinomyces as Actinomyces neuii subsp. neuii* sp. nov., and *Actinomyces neuii subsp. anitratus* subsp. nov. *International Journal of Systematic Bacteriology* **44**: 167-171.
Lawson, P.A., Perez, P.L., Hutson, F.A., Hippe, H. and Collins, M.D. (1993). Towards a phylogeny of clostridia based on 16s rRNA sequences. *FEMS Microbiology Letters* **113**: 87-92.

Funke, G., Stubbs, S., Altwegg, M., Carlotti, A. and Collins, M.D. (1994). *Turicella otitidis* gen. nov., sp. nov., a coryneform bacterium isolated from patients with otitis media. *International Journal of Systematic Bacteriology* **44**: 270-273.
Ash, C. and Collins, M.D. (1992). Comparative analysis of 23S ribosomal RNA gene sequences of *Bacillus anthracis* and emetic *Bacillus cereus* determined by PCR-direct sequencing. *FEMS Microbiol. Lett.* **94**: 75-80.
Cai, J., Collins, M.D., McDonald, V. and Thompson, D.E. (1992). PCR cloning and nucleotide sequence determination of the 18S rRNA genes and internal transcribed spacer 1 of the protozoan parasites *Cryptosporidium parvum* and *C. muris. Biophys. Acta* **1131**: 317-20.
Harland, N.M., Leigh, J.A. and Collins, M.D. (1993). Development of gene probes for the specific identification of *Streptococcus uberis* and *Streptococcus parauberis* based upon large subunit rRNA gene sequences. *J. Appl. Bacteriol.* **74**: 526-531.
East, A.K. and Collins, M.D. (1993). Molecular characterization of DNA encoding 23S rRNA and 16S-23Q rRNA intergenic spacer regions of *Aeromonas hydrophilia. FEMS Microbial. Lett.* **106**: 129-134.
East, A.K., Thompson, D.E. and Collins, M.D. (1992). Analysis of operons encoding 23S rRNA of *Clostridium botulinum* type A. *J. Bacteriol.* **174**: 8158-8162.
Martinez-Murcia, A.J., Harland, N.M. and Collins, M.D. (1993). Phylogenetic analysis of some leuconostocs and related organisms at determined from largesubunit rRNA gene sequences: assessment of congruence of small- and large subunit rRNA derived trees. *J. Appl. Bacteriol.* **74**: 532-541.
Collins, M.D., Lawson, P.A., Willems, A., Cordoba, J.J., Fernandez-Garayzabel, J. Garcia, P., Cai, J., Hippe, H. and Farrow, J.A.E. (1994). The phylogeny of the genus *Clostridium. International Journal of Systematic Bacteriology*. (in press).
Amann, R., Zarda, B., Trebesius, K.H., Ludwig, W. and Schleifer, K.H. (1994). Rapid identification and *in situ* detection of microorganisms using fluorescent rRNA-targeted oligonucleotides, p. 237-244. In: Rapid Methods and Automation in Microbiology and Immunology, R.C. Spencer, E.P. Whright and S.W.B. Newsom (eds.). *Intercept Ltd., Andover, UK.*
Aznar, R., Ludwig, W., Amann, R.I. and Schleifer K.-H. (1994). Sequence determination of rRNA genes of pathogenic *Vibrio* species and whole-cell identification of *Vibrio vulnificus* with rRNA-targetted oligonucleotide probes. *Int. J. system. Bacteriol* **44**: 330-337.
Aznar, R., Ludwig, W. and Schleifer K.-H. (1993). Ribotyping and randomly amplified polymorphic DNA analysis of *Vibrio vulnificus* biotypes. *System. Appl. Microbiol.* **16**: 303-309.
Burggraf, S., Mayer, T., Amann, R, Schadhauser, S., Woese, C.R. and Stetter, K.O. Identifying Archaea with rRNA-targetted oligonucleotide probes. *Appl. Environ. Microbiol.* (in press).
Evers, S., Weizenegger, M., Ludwig, W. Schink, B and Schleifer, K.-H. (1993). The phylogenetic positions of *Pelobacter acetvlenicus* and *Pelobacter propionicus. System. Appl. Microbiol.* **16**: 216-218.
Hahn, D., Amann, R., Zeyer, J. (1994). Oligonucleotide probes for the detection and identification of bacteria, pp. 163-171. In: Improving plant productivity with rhizosphere bacteria. *M.H. Ryder, P.M. Stephens and G.D. Bowen (Eds.), CSIRO, Australia.*
Love, C.A., Pater, B.K.C., Ludwig, W. and Stackebrandt, E. (1993). The phylogenetic position of *Dictyoglomus thermophilum* based on 16S rRNA sequence analysis. *FEMS Microbiology Letters* **107**: 317-320.
Ludwig, W., Dorn, S., Springer, N., Kirchof, G. and Schleifer K.H.. Rapid PCR-based preparation of 23S rRNA-targetted, group-specific polynucleotide probes. *Appl. Environ, Microbiol.* (in press).
Ludwig, W., Kirchhof, G, Klugbauer, N., Weizenegger, M., Betzl, D., Ehrmann, M., Hertel, C, Jigl, S., Tatzel, R. Zitzelsberger, H., Liebl, S., Hochberger, M., Lane, D., Wallnofer, P.R. and Schleifer, K.H. (1992). Complete 23S Ribosomal RNA Sequences of Gram-positive Bacteria with a Low DNA G+C Content. *Syst. Appl. Microbiol.* **15**: 487-501.
Ludwig, W., Mittenhuber, G. and Friedrich, C.G. (1993). Transfer of *Thiosphaera pantotropha to Paracoccus denitrificans. Int. J. System. Bacteriol.* **43**: 363-367.
Manz, W., Szewzyk, u., Erikkson, P. Amann, R., Schleifer K.H. and Stenström, T.A. (1993). *In situ* identification of bacteria in drinking water and adjoining biofilms by hybridization with 16S and 23S rRNA(directed fluorescent oligonucleotide probes. *Appl. Environ. Microbiol.* **59**: 2293-2298.

Roller, C., Wagner, M., Amann, R., Ludwig, W. and Schleifer, K.H. (1994). *In situ* probing of gram-positive bacteria with high DNA G+C content by using 23S rRNA-targeted oligonucleotides. *Microbiol.* (in press).
Rosello-Mora, R., Ludwig, W and Schleifer, K.H. (1993). *Zoogloea ramigera*: a phylogenetically diverse species. *FEMS Microbiol. Lett.* **114**: 129-134.
Sievers, M., Ludwig, W. and Teuber, M. Phylogenetic positions of *Acetobacter, Gluconobacter, Rhodopila* and *Acidiphilium* species as a branch of acidiphilic bacteria in the alpha-subclass of Proteobacteria based on 16S ribosomal DNA sequences. *System. Appl. Microbiol.* (in press).
Spring, S., Amann, R., Ludwig, W., Schleifer, K.H., van Gemerden, H. and Petersen, N. (1993). Dominating role of an unusual magnetotactic bacterium in the microaerophilic zone of a freshwater sediment. *Appl. Environ, Microbiol.* **59**: 2397-2403.
Springer, N., Ludwig, W., Drozanski, V., Amann, R. and Schleifer, K.H. (1992). The phylogenetic status of *Sarcobium lyticum*, an obligate intracellular bacterial parasite of small amoeba. *FEMS Microbiol. Lett.* **96**: 199-202.
Springer, N., Ludwig, W., Amann, R., Schmidt, H.J., Görtz, H.D. and Schleifer, K.H. (1993). Occurence of fragmented 16S rRNA in an obligate bacterial endosymbiont of *Paramecium caudatum. Proc. Natl. Acad. Sci. U.,S.A.* **90**: 9892-9895.
Szewzyk, U., Manz, W., Amann, R., Schleifer, K.H. and Stenström, T.-A. (1994). Growth and *in situ* detection of a pathogenic *Escherichia coli* in biofilms of a heterotrophic water-bacterium by use of 16S- and 23S-rRNA-directed fluorescent oligonucleotide probes. *FEMS Microbiol. Ecol.* **13**: 169-176.
Trebesius, K., Amann, R., Ludwig, W. Mühlegger, K. and Schleifer K.H. (1994). Identification of whole fixed bacterial cells with non-radioactive 23S rRNA-targeted polynucleotide probes. *Appl. Environ, Microbiol.* (in press).
Wagner, M., Amann, R. Kämper, P., Aßmus, B., Hartmann, A., Hutzler, P., Springer, N. and Sclheifer, K.H. (1994). Identification and *in situ* detection of gram-negative filamentous bacteria in activated sludge. *System. Appl. Microbiol.* (in press).
Wagner, M., Amann, R., Lemmer, H. and Schleifer, K.H. (1993). Probing activated sludge with oligonucleotides specific for proteobacteria: inadequacy of culture-dependent methods for describing community structure. *Appl. Environ. Microbiol.* **59**: 1520-1525.
Meunier, J.R. and Grimont, P.A.D. (1993). Factors affecting reproducibility of random amplified polymorphic DNA fingerprinting. *Research in Microbiology* **144**: 373-379.
Grimont, F. and Grimont, P.A.D. (1994). Determination of rRNA gene restriction patterns. In: Howard, J.J. and Walker J.M. 'Diagnostic Bacteriology Protocols' (Methods in Molecular Biology Series). (in press).
De Carvalho, A.F., Gautier, M. and Grimont, F. Identification of dairy *Propiobacterium* species by rRNA restriction patterns. *Research in Microbiology*. (in press).
Brettar, I., Höfle,M.G. (1993). Nitrous oxide producing heterotrophic bacteria from the water colum of the central Baltic: abundance and molecular identification. *Mar. Ecol. Prog. Ser* **94**: 253-265.
Höfle, M.G, Brettar, I. (1994). Taxonomic diversity and metabolic activity of microbial communities in the water column of the central Baltic. *Limnol. Oceanogr.* (submitted).
Katsivela, E., Höfle, M.G. (1994). Separation of transfer RNA and ribosomal RNA using capillary electrophoresis. *J. Chromatogr.* (submitted).
De Vos, P., Van de Woestijne, M., Vancanneyt, M., Verstraete, W. and Kersters, K. (1993). Identification of proferrosamine producing *Pseudomonas* sp. strain GH (LMG 11358) as *Erwinia rhapontici. Syst. Appl. Microbiol.* **17**: 252-255.
Hys, G., Coopman, R., Vancanneyt, M. Kersters, I., Verstraete, W., Kersters, K. and Janssen, P. (1993). High resolution differentiation of aeromonads. *Med. Microbiol. Lett.* **2**: 248-255.
Huys, G., Vancanneyt, M., Coopman, R., Jansen, P., Falsen, E., Altwegg, M. and Kersters, K. (1994). Cellular fatty acid composition as a chemotaxonomic marker for the differentiation of phenospecies and hybridization groups in the genus *Aeromonas. Int. J. Syst. Bacteriol.* (in press).
Huys, G., Kersters, I, Vancanneyt, M., Coopman, R., Janssen, P. and Kersters, K. (1994). Diversity of Aeromonas sp. in Flemish drinking water production plants as determined by gas-liquid chromatographic analysis of cellular fatty acid methyl esters (FAMEs). *J. Appl. Bacteriol.* (submitted).

Kersters, I, Maestrojuan, G.M., Torck, U., Vancanneyt, M., Kersters, K. and Verstraete, W. (1994). Isolation of *Coprothermobacter proteolyticus* from an anaerobic digest and further characterization of the species. *Syst. Appl. Microbiol.* (in press).
Kersters, K., Pot, B., Dewettinck, D., Torck, U, Vancanneyt, M., Vauterin, L. and Vandamme, P. (1994). Identification and typing of bacteria by protein electrophoresis. In: Priest, R.G., A. Ramos-Cormenza and B. Tindail: Identification of bacteria: current status, future prospects. *Plenum Publishing Corporation.* (in press).
Mudarris, M., Austin, B., Segers, P., Vancanneyt, M., Hoste, B., Bernardet J.-F. (1994). Flavobacterium scophtalmum sp. nov., a pathogen of *turbot (Scophthalmus maximus* L). *Int. J. Syst. Bacteriol.* (in press).
Segers, P., Mannheim, W., Vancanneyt, M., De brandt, K., Hinz, K.-H., Kersters, K. and Vandamme, P. (1993). Description of *Riemerella anatipestifer* gen. nov. comb. nov. the causative agent of septicaemia anserum exsudativa and its physiogenetic affiliation within the *Flavobacterium-Cytophaga* rRNA homology group. *Int. J. Syst. Bacteriol.* **43**: 768-776.
Segers, P., Vancanneyt, M., Pot, B., Torck, U., Hoste, B., Dewettinck, D., Falsen, E., Kesters, K. and De Vos, P. (1994). Classification of *Pseudomonas diminuta* Leifson and Hugh 1954 and *Pseudomonas vesicularis* B5sing, D611 and Freytag 1953 in *Brevundimonas* gen. nov. as *Brevundimonas diminuta* comb. nov. and *Brevundimonas vesicularis* comb. nov.*Ant. J. Syst. Bacteriol.* (in press).
Vancanneyt, M. Segers, P., Hauben, L., Hommez, J., Devriese, L.A., Hoste, B., Vandamme, P. and Kersters, K. (1994). *Flavobacterium meningosepticum*, a pathogen in birds. *J. Clin. Microbiol.* (submitted).
Vancanneyt, M. Coopman, R., Tytgat, R., Berny, J.-F., Hennebert, G.L. and Kersters, K. (1992). A taxonomic study of the basidiomycetous yeast *genera Rhodosporidium* and *Rhodotorula* based on whole-cell protein patterns, DNA base compositions and coenzyme Q types. *J. Gen. Appl. Microbiol.* **38**: 363-372.
Vancanneyt, M. Coopman, R., Tytgat, R., Hennebert, G.L. and Kersters, K. (1994). Whole-cell protein patterns, DNA base composition and coenzyme Q types in the yeast genus *Cryptococcus* Kützing and related taxa. *Syst. Appl. Microbiol.* **17**: 65-75.
Vandamme, P., Segers, P., Vancanneyt, M., Van Hove, K., Mutters, R., Hommez, J., Dewhirst, F.E., Paster, B.J., Kersters, K, Falsen, E., Devriese, L.A., Bisgaard, M., Hinz, K.-H. and Mannheim, W. (1994). Description of *Omithobacterium rhinotracheale* gen. nov., sp. nov. isolated from the avian respiratory tract. *Ant. J. Syst. Bacteriol.* **44**: 24-37.
Neefs, J.-M., Van de Peer, Y., De Rijk, P., Chapelle, S., De Wachter, R. Compilation of small ribosomal subunit RNA structures. *Nucl. Acids Res.* **21**: 3025-3049 (1993).
Van de Peer, Y and De Wachter, R. TREECON: A software package for the construction and drawing of evolutionary trees. *Comput. Applic. Biosci.* **9**: 177-182 (1993).
De Rijck, P. and De Wachter, R. DCSE, an interactive tool for sequence alignment and secondary structure research. *Comput. Applic. Biosci.* **9** : 735-740 (1993).
Van de Peer, Y., Neefs, J.-M., De Rijck, P., De Vos, P., De Wachter, R. (1994). About the Order of Divergence of the Major Bacterial Taxa during Evolution. *Syst. Appl. Microbiol.* **17**: 32-38 (1994).

T-PROJECT

'ANIMAL CELL BIOTECHNOLOGY'

Control of recombinant protein glycosylation under defined cultivation conditions (BIOT CT-920304)

COORDINATOR:
H.S. CONRADT, GBF, Braunschweig, D
PARTICIPANT:
A. VERBERT, Lab. Chimie Biologique, Univ. des Sciences & Technologies Lille, Villeneuve d'Ascq Cedex, F

BACKGROUND INFORMATION

The majority of proteins/cytokines destined for human therapy are glycoproteins and require expression and biotechnological production from mammalian host cell lines. Their carbohydrate pattern may effect their pharmaceutical potential. It is important to analyze the specific products obtained under defined production conditions as well as the cellular mechanisms underlying the complex protein glycosylation process.

OBJECTIVES AND PRIMARY APPROACHES

According to the original workplan, in the first year of the project scientific work centered at the comparison of the N- and O-glycosylation of three recombinant model glycoproteins produced under different cell culture conditions. An HPLC-based oligosaccharide-'mapping procedure' for N-glycans and, for analysis at the cellular level, a detailled characterization of the lipid-linked oligosaccharide intermediates of cells cultivated under defined cell culture conditions was planned. We intended to establish a 'reference oligosaccharide bank' for the characterization of glycans of recombinant glycoproteins derived from BHK-21 cells.

The isolation of high producer cell clones and the comparison of recombinant glycoprotein products (with respect to oligosaccharide structure characteristics) from the cell lines obtained with those of parental cell lines was planned.

The results obtained after one year of the contractual period lead us to consider also novel genetically engineered model glycoprotein variants (containing e.g the natural N-glycosylation sites of human EPO and glycosylation domains of human antithrombin III) as possible tools for studying protein N-glycosylation from recombinant mammalian cell lines (see below).

RESULTS AND DISCUSSION

A. Recombinant Glycoprotein Analysis

BHK-21 cells were grown in 2-liter continously perfused stirred bioreactors in suspension as well as on microcarriers both in the presence and absence of fetal calf serum. Furthermore, cells were cultivated in the presence/absence of glutamine, in the presence of ammonia or glucosamine in order to analyse the effect of these medium constituents on the carbohydrate structure of the recombinant product. The purified products were thoroughly characterized by western-blotting, amino acid sequence analysis as well as by carbohydrate mapping using HPAEC-PAD analysis of the PNGase F liberated N-glycans and by mass spectroscopy (LSI-MS, MALDI-MS, NMR) of individual oligosaccharides obtained after anion-exchange chromatography and HPLC-chromatography on NH_2-bonded phase. Thus, proximal α1-6 fucosylation, the sialylation degree as well as the anten-

narity of N-glycans from glycoprotein products obtained under different cultivation conditions have been analyzed:

1. cells grown in suspension secreted a higher proportion of N-glycosylated protein when compared to cells grown on microcarrier;

2. terminal sialylation and proximal α1-6 fucosylation of N-glycans was found to be lower in the product secreted from cells grown in the presence of 2% fetal calf serum than in the product harvested from serum-free culture supernatants;

3. the presence of fetal calf serum in the culture medium affected the degree of O-glycosylation of the model protein;

4. the ratio of biantennary and triantennary N-glycans differed in the product from suspension cultures when compared to that from cells grown on microcarrier both in the presence as well as in the absence of FCS;

5. products from cells grown in glutamine-free medium showed a homogeneous oligosaccharide HPAEC-profile with almost only biantennary fucosylated N-glycans. The presence of glutamine, ammonia or glucosamine in the culture medium resulted in the secretion of glycoproteins with higher-antennary structures;

6. during our studies two BHK cell lines were identified differring in their potential to build up N-glycans on recombinant glycoproteins with terminal GalNAc. Since BHK cells have a low level of a GalNAc-R: sialyltransferase, oligosaccharides from glycoproteins derived from the BHK(A) cells exhibited a lower degree of sialylation state when compared to those isolated from the product from the BHK(B) cells.

B. Analysis of Cellular Parameters

The lipid-linked precursor oligosaccharides and the total cellular glycoprotein bound oligosaccharides was examined under different cell cultivation conditions. In the presence or absence of FCS the same oligosaccharide-PP-Dol pattern was detected in BHK cells under glucose-rich conditions only Glc_3Man_9, Glc_2Man_9 or Glc_1Man_9 were detected; in the glycoprotein fraction Glc_1Man_9 or Man_9 oligosaccharides were predominating.
These patterns were the same for cells grown under suspension or grown on microcarriers. In contrast, when grown under glucose-free conditions, microcarrier attached cells and suspension cells showed a totally different oligosaccharide-PP-Dol pattern with mainly Glc_3Man_9 and Glc_2Man_9 for the suspended cells and Man_2-Man_9 and Glc_1Man_9 for attached cells. Interestingly, the differences detected in the lipid-donors were not observed in the cellular glycoprotein fraction which contained Man_4-Man_9 and Glc_1Man_9 under either conditions.

1. ***Changes of cell culture conditions over a short time period*** (for 1-2 days under perfusion conditions; e.g. omission of amino acids, source of energy supply, presence/absence of FCS) were shown to have no effect on the final glycosylation pattern of the recombinant glycoprotein products.

2. ***Proteolytic degradation:*** a dipeptidyl-peptidase-like activity is presumably responsible for the removal of NH_2-terminal A-P sequence of our model proteins; subsequently, removal of small amino acids (Thr, Ser) leads to truncations of the product and thus increases the heterogeneity of the recombinant proteins.

Similar truncations were also found in the recombinant glycoproteins when expressed from baculovirus-infected insect cells (Sf-21 and Sf-9 cells)

3. ***Comparison of high producer vs. low producer cell clones:*** no differences in the oligosaccharide pattern of the recombinant glycoprotein product was detected in cell several clones that produced about 0.05μg, 1 μg or 10μg of protein/ 24 hours when grown in tissue culture flasks.

4. ***Construction and analysis of BHK-21 cell lines secreting novel variant glycoproteins:*** Cell lines secreting novel genetically engineered glycoprotein variants (protein variants that contain 12-15 amino acid residues around to the N-glycosylation consensus tripeptide sequences of human AT III and human EPO) have been constructed and have been investigated with respect to their carbohydrate structures when expressed from BHK(A) and BHK(B) cells. The results obtained allow for generalizations to be made with respect to the effect of cell culture conditions on the final glycosylation pattern of glycoproteins.

MAJOR SCIENTIFIC BREAKTHROUGHS

The results obtained indicate that under well defined and controlled cell culture conditions (such as applied in the present study by using stirred perfusion bioreactors) different glycoforms of recombinant therapeutic proteins can be obtained with carbohydrate stuctures of predictable physiological properties (the antennarity and sialylation state of oligosaccharides on glycoproteins have been shown to profoundly affect e.g. their in vivo half-lives).

The intracellular amount of precursors for N-glycosylation of proteins is affected by cell culture conditions and therefore is an important parameter for the overall glycosylation characteristics of a final recombinant glycoprotein product. Cell culture conditions can effect the final yield of recombinant protein glycosylation.

The isolation and the detailled chemical characterization of the N-glycan repertoire from recombinant glycoproteins enables the definition of a reference standard for an 'oligosaccharide-mapping' of glycoproteins. This is indispensible for the quality control of recombinant glycoprotein products from a given host cell line destined for clinical use (*lot-to-lot consistency*).

General considerations

Methods for cultivation of recombinant mammalian cell lines for the production of therapeutic glycoproteins have been evaluated with respect to the final glycosylation pattern of a desired product. Investigations included: at the level of characterisation of the final product *the polypeptide and carbohydrate structural analysis of glycoproteins and at the cellular level a detailled analysis of physiological parameters such as concentrations and turnover of oligosaccharide precursors* that can be expected to affect the glycoforms/glycan structures of products when using different cultivation conditions. A final aim of this project was to enable the 'modelling' of animal cell culture to achieve the reproducible production of defined glycoprotein using a given cell line and furthermore, to allow for predictions to be made with respect to final obtain glycoform-products of the expected biological properties. The results indicate that the quality of products is only marginally affected by short-term changes of cell culture conditions. Long-term alterations in cell culture can be designed generating glycoprotein products with predictable glycan structures and hence with defined physiological properties. This can be achieved, e.g. using cell culture conditions in the absence or presence of glutamine/ammonia or cultivating cells on microcarriers or as suspension.

MAJOR COOPERATIVE LINKS

Two members of the Lille group visited the GBF. Experiments have been carried out together with members of the GBF group (two weeks). Recombinant as well as wild-type host cells have been exchanged.
A member of the GBF group visited the Laboratoire de Chimie Biologique for the exchange of results and discussion of further activities.
Glycosylation analysis (GBF group) was carried out for one laboratory participating in the BRIDGE programme.
Standard oligosaccharides for 'mapping of carbohydrates' have been made available to other laboratories participating in the BRIDGE programme.

PUBLICATIONS

Joint publications

Cacan, R., Labiau, O, Mir, A.-M. and Verbert, A. (1993). Effect of cell attachment and growth on the synthesis and fate of dolichol-linked oligosaccharides in Chinese hamster ovary cells. *Eur. J. Biochem.* **215**, 873-881

Gawlitzek M., Villers C., Verbert A., Wagner R. and Conradt H.S. (1994). Changes in the glycosylation pattern of recombinant proteins effected by defined culture conditions of BHK-21 cells. In: *Animal Cell Technology: Products of Today, Prospects of Tomorrow*, Spier, R.E., Griffiths, J.B. and Berthold, W. (eds.) Butterworth -Heinemann, Oxford, UK. pp. 649-651

Hermentin P., Donges R., Witzel R., Hokke C.H., Vliegenthart J.F.G., Kammerling J., Conradt H.S., Nimtz M. and Brazel D. (1994). A strategy for the mapping of N-glycans by high-performance capillary electrophoresis. *Anal. Biochem.*, (in press)

Hoffmann, A., Nimtz, M., Wurster, U. and Conradt, H.S. (1994) Carbohydrate structures of β-trace protein from human cerebrospinal fluid: Evidence for 'brain-type' N-glycosylation; *J. Neurochem.*, (in press)

Lütkemeyer D., Riese U., Heidemann R., Bretschneider M., Lehmann J. and Conradt H.S. (1994). Product quality studies of antithrombin III during medium recycling in pilot scale. In: *Animal Cell Technology: Products of Today, Prospects of Tomorrow*, Spier, R.E., Griffiths, J.B. and Berthold, W. (eds.) Butterworth -Heinemann, Oxford, UK. pp. 760-763

Villers, C., Plancke, Y., Cacan, R. and Verbert, A. (1994) Rapid detection of the alternative N-glycosylation pathway using high pH anion exchange. *Glycoconjugate J.*, (in press)

Villers, C., Cacan, R., Mir, A.-M., Labiau, O. and Verbert, A. (1994). Release of oligomannoside-type glycans as a marker of the degradation of newly synthesized glycoproteins. *Biochem. J.*, **298**, 135-142

Individual publications

Björk, J., Ylinenjärvi, K., Olson, S.T., Zettlmeissl, G., Conradt, H.S. and Hermentin, P. (1992) Decreased activity of recombinant ATIII for Heparin due to increased glycosylation, *Biochem. J.*, **286**, 793-800

Gawlitzek, M., Karger, C., Conradt. H.S. and Wagner, R. (1992). Influence of diffrent cultivation conditions on the N-glycosylation pattern of recombinant pharmaproteins. *Dechema Biotechnology Conferences*, vol. 5 part A, pp. 281-285

Grabenhorst E., Karger C., Reis U., Collins J., Lindenmaier W., Rollwage K., Jäger V. and Conradt H.S. (1992). Processing of Recombinant Proteins: Comparison of N- and O-Glycosylation of Recombinant Human Glycoproteins in Heterologous Mammalian Cells and Baculovirus-Infected Sf 21 Cells. In: *Baculovirus and Recombinant Protein Production Processes*, Editiones Roche, Basel, Switzerland 1992, pp. 145-156; Vlak, J.M., Schlaeger, E.-J. and Bernard, A.R. eds.

Grabenhorst E., Hofer B., Nimtz M., Jäger V. and Conradt H.S. (1993). Biosynthesis and secretion of human interleukin-2 glycoprotein variants from baculovirus infected Sf21 cells: characterization of polypeptides and posttranslational modifications. *Eur. J. Biochem.*, **215**, 189-197

Hermentin, P., , Nimtz, M., Seidat, M., Witzel, M., Kammerling, H., Vliegenthart, J.H.F. and Conradt, H.S. (1992). A strategy for the mapping of N-glycans by High-pH Anion-Exchange Chromatography with pulsed Amperometric Detection. *Anal. Biochem.*, **203**, 281-289

Hoffmann A., Conradt H.S., Gross G., Nimtz M., Lottspeich F. and Wurster U. (1992). Purification and Chemical Characterization of ß-trace Protein from Human Cerebrospinal Fluid: its Identification as Prostaglandin D Synthase. *J. Neurochem.*, **61**, 451-156
Nimtz, M., Wray, V., Augustin, M., Klöppel, K.-D., Conradt, H.S. (1993). Native Oligosaccharide Structures of Human Erythropoietin Expressed From BHK-21 Cells. *Eur. J. Biochem*, **213**, 39-56

Optimization and validation of virus-based linear vectors (BIOT CT-910305)

COORDINATOR:
J. ROMMELAERE, ULB, Brussels, B
PARTICIPANT:
S. RIVA, IGBE-CNR, Pavia, I

BACKGROUND INFORMATION

Parvoviruses posses some properties that render them particularly interesting as eukaryotic expression vectors. Autonomous parvoviruses replicate as episomes in many different cell types. Replication starts from one of the palindromic extremities of the genome. Relatively high levels of protein expression are obtained in permissive cells.

The genome of the autonomous parvovirus MVM (Minute Virus of Mice) has been cloned. When this molecular clone is transfected into permissive cells, the viral DNA is excised from the pBR322 backbone and replicates autonomously as a virus. MVM transcription occurs from two promoters: P4, the early promoter, controls the expression of two overlapping genes coding for non-structural proteins; the non-structural NS-1 protein plays a role in the cytotoxicity and the replication of the parvovirus and it transactivates the second promoter, P38, that controls the expression of two capsid proteins (VP1 and VP2).

MVM is a lytic virus but the killing effect is mainly restricted to transformed cells. The difference in susceptibility between normal and transformed cells is regulated at the level of viral DNA amplification and/or gene expression. The tropism of parvoviruses for transformed cells can be exploited in cancer therapy to target the expression, to tumour cells, of genes that code for toxins or proteins that can locally stimulate an anti-tumour immune response.
MVM is an interesting basis for the construction of a linear vector for mammalian cells. Its DNA replicates in the nucleus as a linear extrachromosomal molecule. A linear construct consisting of a chromosomal origin of DNA replication linked to the MVM telomeric termini could constitute an innovative vector that replicates autonomously in the nucleus of mammalian cells. Such a vector, being subject to physiological cell cycle controls, could be useful, for example in gene therapy.

OBJECTIVES AND PRIMARY APPROACHES

As a model system we have chosen to clone the human IL2 gene under the control of the P38 promoter. Recombinant IL2 is already used in cancer therapy but the high doses needed provoke many side effects. Furthermore, frequent injections are necessary, due to the short half-life of recombinant IL2. Problems encountered with currently used lymphokine therapy could be alleviated by a more efficient delivery system such as the parvoviral vector proposed here. In this vector 1.5 kb of capsid genes have been replaced by the 500 base-pair cDNA of human IL2. When transformed cells are transfected with the plasmid carrying this vector, they secrete IL2 and they produce recombinant (MVM/IL2) viral particles when they are cotransfected with a helper plasmid that provides VP genes in trans. The titres of MVM/IL2 stocks originally obtained in this way were low, about 5.10^3 particles/ml, and they were contaminated by wild-type parvovirus which arises through

recombination between the plasmids carrying the MVM/IL2 genome or the VP genes.

Our main objectives were to

(i) increase the yield of MVM/IL2 and to produce stocks that are free of wild-type MVM,
(ii) to set up an animal model in order to evaluate, *in vivo*, the anti-tumour potential of the recombinant vector.

Another aim was the development of a linear vector for mammalian cells based on parvovirus telomeric ends and on a replication origin isolated from the human cell DNA. Thanks to the cellular replication origin, such a vector should be subject to physiological cell-cycle regulation mechanisms.

The decision to develop a linear vector instead of the more traditional circular ones derives from the observation that it has been impossible so far to observe any autonomous replication of small circular DNA molecules in mammalian cells even when they contain putative cellular replication origin such as that associated to the DHFR gene. These failures can be due either to topological constrains met during the replication of small circular DNA molecules or to the lack of appropriate nuclear retention signals. The use of vectors based on parvoviral DNA could overcome these problems.

Two parallel approaches were used:

a) isolation and cloning of an active human origin of DNA replication
b) study of the mode of replication of MVM DNA containing an origin of bidirectional replication (viral or cellular) cloned into the capsid protein gene.

RESULTS AND DISCUSSION

Transfection of permissive cells with a plasmid that carries the wild-type MVM genome produces about 10 fold more virus than a similar transfection with a pMVM/IL2 plasmid. We could show that packaging and not replication of the MVM/IL2 DNA is limiting in the production of infectious particles. We have therefore constructed a second version of the MVM/IL2 vector in which 1kb of VP coding sequences have been added back again, thus restoring a DNA structure and size that are closer to those of wild-type MVM. With this vector 10 fold more infectious MVM/IL2 was produced.

The above described cotransfections of a pMVM/IL2 plasmid with a helper plasmid (carrying VP genes) were performed in NB-E cells, a human new born kidney cell line. A clone was derived from these cells which has integrated a modified MVM genome. These cells express VP genes under the control of the natural P38 promoter whereas the expression of NS proteins is controlled by an inducible promoter, the MMTV LTR. Only non-toxic amounts of NS-1 are produced in the uninduced state but induction of the MMTV LTR promoter leads to cell death. When these cells are transfected with the pMVM/IL2 plasmid, infectious recombinant virus particles are produced. These experiments have shown that packaging of viral DNA can be obtained with the same efficiency whether VP proteins are produced from integrated genes or from genes carried on a helper plasmid. It should be possible to establish a packaging cell line from cells that produce more efficiently infectious virus after transfection than NBE cells. We have shown that this is the case for the simian cos-1 cells which produce about 10 fold more MVM/IL2 after cotransfection than NB-E cells. However attempts to isolate a similar packaging cell line from these cells have failed, maybe because

they are sensitive to the small amount of NS-1 produced from the uninduced MMTV LTR. We are currently investigating the capacity of other cell lines to produce infectious virus after transfection of a molecular clone of parvovirus DNA.

In parallel with the optimalization of MVM/IL2 virus production, different murine tumour cell lines, for which a syngeneic host is available, have been tested as to their permissiveness (expression of viral genes and sensitivity) to MVM *in vitro*.

Candidate tumour cell lines are:

(i) B16, a C57 black mouse melanoma,
(ii) P815, a mastocytoma cultured from a DBA/2 mouse tumour induced by methylcolanthrene,
(iii) SA1, an Ajax mouse fibrosarcoma,
(iv) FS29, a methyl chlorine induced C57 black mouse mammary carcinoma and
(v) A9HT, a metastatic variant of A9 cells which are the classical mouse cell line used for MVM virus production. A9 are derived from C3H/an mouse L cells.

Only A9, A9HT and P815 cells expressed consistenly, significant amounts of IL2 after infection with MVM/IL2 recombinant virus. These cells are also sensitive to the killing effect of MVM(p); the survival to a multiplicity of infection of 5 viral particles/cell was 20% to 30% for P815 or A9 and A9HT respectively. Moreover, since P815 cells give rise to tumours about 10 days after subcutaneous injection in DBA/2 mice, they will be used for the in vivo studies on the anti-tumour effect of MVM/IL2. Wild-type MVM or recombinant MVM/IL2 infected tumour cells will be implanted into syngeneic mice and tumour take will be compared to that achieved with non infected cells. If a protective effect is observed with MVM/IL2, wild-type MVM or MVM/IL2 will be inoculated into established tumours to evaluate the contribution of the IL2 induced immunomodulation in tumour rejection.

The isolation of a chromosomal origin of replication was fully successfull. A human sequence containing a bidirectional origin was cloned and extensively characterized. It was found to map on ch. 19 p 13.3 and to be located in 500 nt between the 3'-end of the gene for the lamin B2 and the 5'-end of another unidentified gene. The origin overlaps an active promoter and it contains several binding sites for protein factors. The preliminary work on MVM linear constructs containing a viral origin of DNA replication (SV40 or EBV) evidenced a number of unexpected problems. In detail, when the SV40 replication origin is cloned in the capsid protein gene, the parvoviral mode of DNA replication (from the terminal palindrome) is predominant while the SV40 bidirectional model is inhibited in cos-1 transfected cells. On the other hand when a similar construct with non functional NS-1 protein is tested for autonomous replication in Cos-1 cells, only the replication of circular molecules is observed. Thus the NS-1 protein, which is required for the replication of linear termini, is inhibitory for the viral origins (and probably also for a cellular origin). Obviously this inhibitory effect must be eliminated, before the originally planned construct containing the cellular replication origin can be tested. Experiments in this direction are in progress.

MAJOR SCIENTIFIC BREAKTHROUGHS

We have succeeded in increasing about 100 fold the titres of recombinant virus obtained by cotransfecting cos-1 cells with a capsid protein-producing helper plasmid and a modified MVM/IL2 vector the size of which is close to that of wild-type MVM. We are now focusing on the construction of a packaging cell line to

produce wild-type MVM-free stocks of MVM/IL2 and on methods to concentrate these stocks for in vivo infections.
We have succeeded in isolating and cloning a 'bona fide' human DNA replication origin which is now available for vector construction. Moreover a method for the fine mapping of replication origins based on competitive PCR, was developed.

We have demonstrated that the MVM non structural protein NS-1 trans-inhibits viral origins of replication (SV40 and EBV) and possibly also genomic replication. On the other hand we demonstrated that the insertion of additional sequences (> 3000 nt) in the MVM genome inhibits replication of linear DNA molecules from viral termini in transfected mammalian cells. Both pieces of preliminary information are essential for further development of MVM based linear vectors.

MAJOR COOPERATIVE LINKS

Cooperation was enforced through the exchange of materials and of technical information. Joint meetings were held in the course of the project.

PUBLICATIONS

Joint publication
L. Tenenbaum, F. Dupont, P. Spegelaere, L. Zentilin, P. Norio, M. Giacca, S. Riva, A. Falaschi and J. Rommelaere — Inhibition of heterologous DNA replication by the MVMp nonstructural NS-1 protein: identification of target sequence. Virology **197**, 630-641 (1993)

Individual publications
G. Biamonti, G. Perini, F. Weighardt, S. Riva, M. Giacca, P. Norio, L. Zentilin, S. Diviacco, D. Dimitrova and A. Falaschi — A human DNA replication origin: localization and transcriptional characterization. Chromosome **102**, S24-S31 (1992).
G. Biamonti, M. Giacca, G. Perini, G. Contreas, L. Zentilin, F. Weighardt, M. Guerra, G. Della Valle, S. Saccone, S. Riva and A. Falaschi — The gene for a novel human lamin maps at a highly transcribed locus of chromosome 19 which replicates at the onset of S-phase. Molec. Cell. Biol. **12**, 3499-3506 (1992)
M. Giacca, L. Zentilin, P. Norio, S. Diviacco, D. Dimitrova, G. Contreas, G. Biamonti, G. Perini, F. Weighardt, S. Riva and A. Falaschi — Fine mapping of a replication origin of human DNA. Proc. Natl. Acad. Sci. USA, 1994 (in press).
S.J. Russell, A. Brandenburger, C.L. Flemming, M.K.L. Collins and J. Rommelaere. Transformation-dependent expression of interleukin genes delivered by a recombinant parvovirus. J. Virol. **66**, 2821-2828 (1992).

Improvement of production of bioactive proteins by genetically engineered animal cells using the novel furin-class of mammalian endoproteolytic processing enzymes (BIOT CT-910302)

COORDINATOR:
W.J.M. VAN DE VEN, CME, KU Leuven, Leuven, B
PARTICIPANTS:
J. BULLERDIEK, University of Bremen, Bremen, D
J. HUTTON, University of Cambridge, Cambridge, GB
G. KEIZER, Holland biotechnology bv, Leiden, NL
L. LANFRANCONE, University of Perugia, Perugia, I
G. MARTENS, University of Nijmegen, Nijmegen, NL
J. van MOURIK, CLB, Amsterdam, NL

BACKGROUND INFORMATION

An increasing number of therapeutic products from the Biotechnology sector are now being produced in animal cells. The aim of this project was to increase production of such compounds and to provide a greater understanding of the underlying scientific principles which result in their expression in animal cells.

OBJECTIVES AND PRIMARY APPROACHES

1. K.U. Leuven: Proprotein processing using the novel furin-class of processing enzymes.
2. University of Cambridge: Expression of active forms of the furin-like prohormone converting endopeptidases.
3. CLB Amsterdam: Processing of precursor proteins for blood clotting factors.
4. University of Nijmegen: Identification of genes involved in sorting, post-translational modification and secretion of peptide hormones.
5. University of Perugia: Expression of cloned genes into mammalian cells using retroviral- and EBV-derived vectors.
6. University of Bremen: Development of extrachromosomal vectors for introduction and amplification of foreign genes in animal cells.
7. Holland biotechnology bv: Production of bioactive proteins by animal cells in large scale cell culture.

RESULTS AND DISCUSSION

A. K.U. Leuven: Proprotein processing using the novel furin-class of processing enzymes.

1. Structure and function of eukaryotic proprotein processing enzymes of the subtilisin family of serine proteases

Production of a broad spectrum of regulatory proteins in eukaryotes occurs via an intricate cascade of biosynthetic and secretory processes. Often, these proteins are initially synthesized as parts of higher molecular weight, but inactive, precursor proteins. Specific endoproteolytic processing of these proproteins is required to generate the regulatory proteins in a mature and biologically active form. The first known eukaryotic enzyme with the exquisite cleavage specificity for paired basic amino acid residues was the prohormone processing enzyme kexin [EC 3.4.21.61],

a subtilisin-like serine protease which is encoded by the KEX2 gene of yeast *Saccharomyces cerevisiae*. Recently, a number of kexin-like mammalian proprotein processing enzymes were discovered. The enzyme furin, which is encoded by the FUR gene, was the first and can be considered as the prototype of a mammalian subclass of subtilisin-like serine proteases. It is predicted to contain a 'prepro' domain, a subtilisin-like catalytic domain, a middle domain, a cysteine-rich region, a transmembrane anchor and a cytoplasmic domain. Furin is expressed in a wide variety of tissues, if not in all. In all likelihood, it is the enzyme responsible for the proteolytic bioactivation of a wide variety of precursor proteins.

2. Generation of structural and functional diversity in furin-like proteins in Drosophila melanogaster by alternative splicing of the DFUR1 gene

DNA sequences encoding kexin- and furin-like candidate proprotein processing enzymes have also been identified in *Drosophila melanogaster*; the corresponding genes have been designated *DFUR*1 and *DFUR*2. Expression of the *DFUR*1 gene apparently results in the production of three candidate proprotein processing enzymes; each has the same subtilisin-like catalytic domain but there are major differences in other regions of the proteins. This diversity is generated by alternative splicing. The *DFUR*2 gene product, Dfurin2, has a very large cysteine-rich region and possesses a similar proprotein processing activity as mammalian furin.

3. Subtilisin-like proprotein convertases: Homology modelling of the catalytic domain of human furin

A model is presented for the three-dimensional structure of the catalytic domain of the human serine proteinase furin, and its interaction with model substrates, based on the crystal structures of subtilisin BPN' and thermitase in complex with the inhibitor eglin. Based on this model, protein engineering can be employed not only to test the predicted enzyme-substrate interactions, as demonstrated for human furin, but equally importantly to design proprotein convertases with a desired specificity, or to design novel substrates or inhibitors.

4. Modulation of furin-mediated proprotein processing activity by site-directed mutagenesis

Our studies indicate that the three residues of the catalytic triad of furin, Asp46, His87 and Ser261, are critical not only for substrate processing but also for maturation of furin. Furthermore, evidence is provided that maturation of furin occurs through an intramolecular autocatalytic process. Analysis of carboxy-terminal deletion mutants revealed that the segment encompassing residues E449 to E469 of the 'middle' domain, which is more than 100 residues downstream of the predicted catalytic domain, contains residues which seem to be critical for processing activity but that the more carboxy-terminal cysteine-rich region, the transmembrane region and the cytosolic tail are dispensable. Finally, it was demonstrated that particular negatively charged residues in or near the substrate-binding region of furin are critical for cleavage activity and specificity of the enzyme for multiple basic residues in the substrate.

5. Expression of the dibasic proprotein processing enzyme furin is directed by multiple promoters

The prototype mammalian proprotein processing enzyme furin is shown to be encoded by three distinct *FUR* mRNA isoforms which differ only in their 5'-untranslated regions. By primer extension analysis, the transcription start sites of the three mRNA isoforms were defined. The genomic regions located immediately upstream of the three alternative transcriptional start sites were shown to possess

promoter activity in transfection experiments using the luciferase encoding gene as reporter. In a liver cell line, the P1 promoter appeared to be the strongest; in a lung cell line, the P1A promoter. Human FUR promoter P1 but not P1A or P1B was transactivated by transcription factor C/EBPβ. Other members of this family of bZIP transcription factors, C/EBPα and C/EBPβ, were not able to transactivate the P1 promoter. Promoter P1A and P1B have characteristics of promoters of house-keeping genes. They lack TATA or CAAT boxes upstream of the transcription start site but are very GC-rich and contain several SP1 sites. Promoter P1, on the other hand, has a TATA box in the proximal promoter region. In electromobility shift assays and DNase I footprinting analysis, transcription factor SP1 was found to bind to the proximal region of the P1 promoter. Altogether, our results indicate that expression of the human *FUR* gene is directed by alternative promoters, house-keeping (GC-rich) as well as regulated (TATA containing) promoters, suggesting that their differential use may be a mechanism to modulate levels of the furin enzyme.

B. University of Cambridge: Expression of active forms of the furin-like prohormone converting endopeptidases

1. Generation of immunological reagents

High titre sera directed at the catalytic domains of PC1 and PC2 were obtained that proved useful for immunofluorescence microscopy, immunoelectron microscopy and immunoprecipitation of radiolabelled proteins. Antibodies which were reactive with active enzyme were obtained only in one instance, however these were of key importance in defining the relationship between PC2 and PC1 and native forms of proinsulin converting activity present in insulinoma secretory granules.

2. Rapid assays and kinetic measurements for PC2 and PC1

An assay system was developed for the rapid assay of these enzymes based upon proinsulin as a substrate and using a monoclonal antibody generated to a peptide spanning the cleavage site in proinsulin for separation of the product and substrate for the reaction. This assay permitted the determination of the kinetic constants for these enzymes for the first time and provided a simple, quantitative, and robust assay for determining the levels of expression and kinetic properties of genetically engineered forms of the enzymes.

3. Expression of PC2 and PC1 in heterologous cells

The expression of active forms of the PC2 and PC1 enzymes from different cell types was confronted by problems relating to the stability of the expressed product in saline media and to a requirement for the intracellular activation of the enzymes. The former problem was addressed by examining the secretion and activity of PC1 and PC2 expressed from cRNA's micro-injected into *Xenopus* oocytes from which kinetic data was obtained for the first time. Optimal conditions were defined which should be applicable to situation where large scale production is required.

4. Post-translational modification of PC2

The post-translational fate of PC2 was followed in pancreatic islet tissue using a combination of pulse-chase radiolabelling experiments and microsequence analysis of the mature protein. It was concluded that the protein is retained in the endoplasmic reticulum for a considerable length of time (2h). During this time it appears to be converted into a shorter form probably by an autocatalytic process.

After cleavage it is sorted into the secretory granule where it undergoes further proteolytic modification which may be important either to its activation or inactivation.

5. Cell biology of the processing enzymes

Constructs have been made of the PC and furin processing enzymes with the view to examining their targetting to intracellular organelles and maturation in cells with a regulated pathway of secretion. Initial transfection experiments have been performed without much success as levels have been low relative to endogenous expression of the enzymes. A new series of constructs are now being made using vectors provided by the Hamburg group which have amplification sequences within them. At the same time epitope tags are being introduced into the molecules.

C. CLB Amsterdam: Processing of precursor proteins for blood clotting factors

1. Endoproteolytic processing of Factor VIII

a) Tissue-specific endoproteolytic cleavage of Factor VIII
In cultured primary cells (smooth muscle cells) Factor VIII-del(868-1652) was predominantly produced in its single chain form. Similarly, in a number of cell lines, including C127, CV1 and COS B cells, this mutant was secreted as a unprocessed, single chain molecule. Likewise the Factor VIII mutant that lacks both putative cleavage sites was produced as a single-chain polypeptide. These observations suggest that the endoproteolytic cleavage after Arg^{1648} is a tissue-specific event.

b) Co-expression of Factor VIII and furin (-like enzymes)
Co-expression with the *FUR* gene and other genes structurally related to the subtilisin family (PC1, PC2, PACE4) does not substantially overcome the apparent limitations in endo-proteolytic cleavage of wild-type Factor VIII or Factor VIII-del(868-1652). These observations suggest that Factor VIII is not a major substrate for furin and related proteases.

2. Effect of endo-propteolytic cleavage on Factor VIII function

a) Interaction of single-chain Factor VIII with Factor IX
Proteolytic cleavage of single-chain Factor VIII by thrombin, a process that accompanies Factor VIII activation, enhances the binding of a monoclonal antibody directed to the Factor VIII polypeptide region involved in Factor IX binding. Thrombin-mediated activation does not affect antibody binding to wild-type, fully processed Factor VIII. As the Factor VIII/IX assembly is essential for the expression of Factor VIII co-factor function, these observations suggest that endoproteolytic processing serves a role in controlling proper exposure of one of the major functional domains on the Factor VIII molecule.

b) Interaction of single-chain Factor VIII with von Willebrand Factor
The affinity ($K_d \approx 10^{-10}$M) of single-chain Factor VIII for von Willebrand Factor does not differ from the affinity of wild-type Factor VIII for its natural carrier protein. Although the primary binding site of von Willebrand Factor (Tyr^{1680}) is located in the vicinity of the major endoproteolytic cleavage sites (Arg^{1648}), cleavage after the latter site apparently does not effect the Factor VIII-von Willebrand Factor interaction.

D. University of Nijmegen: Identification of genes involved in sorting, post-translational modification and secretion of peptide hormones

1. Isolation and characterization of cDNAs encoding novel proteins functioning in the regulated secretory pathway: Construction and differential screening of a* Xenopus *intermediate pituitary cDNA library

The differential hybridization approach revealed that about 75% of the cDNA clones code for the prohormone proopiomelanocortin (POMC). Differentially hybridizing, non-POMC cDNA clones were found to encode known proteins which are present in the secretory pathway, including binding protein BiP, the folding enzyme PDI, the ER calcium-binding protein calreticulin, the clathrin-associated protein AP47, the prohormone convertase PC2, the amidating enzyme PAM, carboxypeptidase E/H, and the neuroendocrine-specific proteins 7B2 and secretogranins 1, 2 and 3.

Three differentially hybridizing, novel cDNAs have been isolated: 1) they encode novel proteins (not present in the DNA or protein sequence data bases), 2) they are only expressed in neurons and endocrine cells (neuroendocrine-specific cDNAs) and 3) they are highly induced in intermediate pituitary cells of black compared to white animals (10- to 15-fold induction).

Full-length clones of two of these novel cDNAs have been isolated and sequenced. Recombinant proteins have been produced and antibodies are currently being made.

2. Cloning and expression of the* Xenopus *prohormone convertase PC2

A *Xenopus* homologue of the prohormone convertase PC2 has been cloned and sequenced. Expression of *Xenopus* PC2 mRNA has been studied with Northern blot analysis, RNase protection assays and in situ hybridization. Biosynthesis of PC2 protein in *Xenopus* intermediate pituitary cells has been examined by pulse-chase analysis in combination with immunoprecipitation experiments. *Xenopus* PC2 is initially synthesized as a 75 kDa proPC2 which is subsequently processed through a 71 kDa intermediate to the mature 69 kDa PC2. Only the 69 kDa form of PC2 was found to be released into the incubation medium. The time course of conversion of the inactive pro-form of PC2 into the mature active PC2 paralleled the time course of POMC and 7B2 biosynthesis and processing in the *Xenopus* intermediate pituitary cells.

3. The neuroendocrine polypeptide 7B2

A major breakthrough of our research within the BRIDGE program involved our discovery of the function of the neuroendocrine polypeptide 7B2. The 7B2 protein has been isolated more than ten years ago but its function has been elusive ever since. The precursor from of newly synthesized 7B2 was found to be co-immunoprecipitated with PC2, while the pro-form of PC2 was cross-linked and immunoprecipitated with 7B2. Therefore, the 7B2 precursor protein interacts with proPC2. In the secretory pathway of *Xenopus* intermediate pituitary cells, this interaction is transient as shown by pulse-chase analysis; association commences early in the endoplasmic reticulum while dissociation in the later stages of the secretory pathway coincides with the processing of POMC, 7B2 and proPC2. In an *in vitro* PC2 enzyme assay, the precursor form of recombinant 7B2 was found to be a potent inhibitor of PC2 ($K_i \sim 6$ nM). The carboxy-terminally processed form of recombinant 7B2 was virtually inactive. Neither form of recombinant 7B2 inhibited PC1 enzyme activity. The precursor form of recombinant 7B2 but not processed 7B2 prevented *in vitro* cleavage of proPC2. We conclude that 7B2 is a

molecular chaperon preventing premature proPC2 activation in the regulated secretory pathway. The 7B2 protein thus constitutes an important component of the mechanisms involved in the production of peptide hormones and neuropeptides by regulating prohormone convertase activity in the secretory pathway.

E. University of Perugia: Expression of cloned genes into mammalian cells using retroviral- and EBV-derived vectors

1. Formation of Shc-Grb2 complexes is necessary to induce neoplastic transformation by overexpression of Shc proteins

We have demonstrated that, *in vitro*, the high affinity binding of Grb2 to Shc proteins requires phosphorylation of Shc at Tyr317, which lies within the high affinity binding motif for the Grb2 SH2 domain, pYVNV, where Asn at the +2 position is crucial for complex formation. *In vivo*, Tyr317 is the major, but not the only, site for Shc phosphorylation, and is the sole Shc high affinity binding site for Grb2. Mutant Shc proteins with substitution of the Tyr317 by Phe lose the capacity to be highly phosphorylated on tyrosine upon growth factor receptor activation, to bind Grb2 and to induce neoplastic transformation. In contrast, Shc proteins that have an extensive aminoterminal deletion, but retain the Tyr317 site and the SH2 domain conserve the capacity to be phosphorylated, to bind to Grb2 and to induce cell transformation. These data indicate that the formation of the Shc-Grb2 complex is a crucial event in the transformation induced by overexpression of Shc and support the notion that Shc proteins can deliver activation signals to RAS.

2. Shc is constitutively phosphorylated in human tumors

We measured the level of constitutive tyrosine phosphorylation in a series of spontaneous human tumours and found that:

a) Shc proteins are constitutively phosphorylated in those tumours with genetic alterations of receptor or cytoplasmic tyrosine kinases. In those tumours, Shc proteins are constitutively associated with the implicated tyrosine kinase and the Sem5/Grb2 product.

b) Shc proteins are constitutively phosphorylated in a tumour subset not previously known to carry genetic alterations of tyrosine kinases. In those tumours Shc proteins are constitutively associated with a phosphotyrosine-containing peptide and the Sem5/Grb2 product.

3. Shc is constitutively phosphorylated in human leukemias and recruit a phosphorylated p140 protein involved in the GM-SCF signalling

In approximately 70% of human leukemias Shc is constitutively phosphorylated and associated with an as yet not identified 140 kDa phosphoprotein (p140). p140 is involved in the GM-CSF signalling pathway as shown by the fact that the association with Shc is induced by GM-CSF in haematopoietic myeloid cell lines.

F. University of Bremen: Development of extrachromosomal vectors for introduction and amplification of foreign genes in animal cells

Today cloning of smaller genomic fragments are standardised methods. However, there are no tools to handle large DNA-fragments in size ranging between whole chromosomes and yeast artificial chromosomes. Furthermore there are no cloning vectors for precise and continuous expression of large genes in mammalian cells. Common tools to investigate these problems are viral circular vectors and linear artificial 'mini chromosomes'.

The objective of our project was to develop a new tool for manipulating large fragments of DNA. We planned to use naturally occurring or artificially constructed extrachromosomal elements, so called Double Minutes (DM), as eukaryotic vectors. The size of DMs varies from about 200 kbp to more than 10 Mbp. They are lacking centromeres, most of them do not have telomeric repeats, they are autonomously replicating and generally circular elements which as a rule are not 1: 1 distributed to the daughter-cell during mitosis. Though their integration into chromosomes seems possible, this is a rare event and DMs are otherwise relatively stable. DMs are naturally found in tumour cells corresponding to amplified genes. If they have a dominant marker gene, their copy number can be increased. The aim of our work was to construct DMs as selectable extrachromosomal cloning vector with an efficient cloning site. The following results were obtained:

1. Establishment of a recombinant DNA-fragment with a dominant selection marker and a rare cutter restriction site.
2. Insert of a cloning site in naturally occurring DMs.
3. Generation of artificial DMs possessing a I-Sce I cloning site with a new enzyme — fusion -hybridization (EFH) technique.

G. Holland biotechnology bv: Production of bioactive proteins by animal cells in large scale cell culture

The main objective was the large scale production of furin in animal cells using a hollow fibre culture system. There are many different types of hollow fibre culture systems on the market all having their own specific advantages and drawbacks. We decided to combine all the different advantages as much as possible by constructing our own system rather than using the hollow fibres which are specifically sold for the use in cell culture.

First the system was tested, using an antibody producing hybridoma, whether the cells should be grown at the inside or at the outside of the fibres. The performance characteristics such as cell growth and immunoglobulin production did not significantly differ between the two systems for the culture period tested. For practical and technical reasons it was chosen to culture the cells at the outside of the fibres.

Two brands (Organon Technika (OT) and Fresenius (F)) of artificial kidneys were selected and tested. The Fresenius artificial kidney had the best performance characteristics. The windings of the fibres was more loose than the winding of the fibres of the OT artificial kidney thereby allowing a better cell growth between the fibres. In addition the material used binds endotoxins and prevents bacterial infections to a large extend.

Different parameters have been investigated:

1. *The amount of cells necessary to start*
2. *The pH versus CO_2 concentration*
3. *The amount of Fetal Calf Serum (FCS) within and outside the fibres*
4. *The immunoglobulin production in the course of time.*

Conclusions:

The results clearly indicate that the use of an inexpensive hollow fibre system (artificial kidney) has a strong potential for upscaling of protein production. In addition after the start-up phase a minimum amount of handling time is necessary to keep the system running thereby allowing running several systems at the same time.

MAJOR SCIENTIFIC BREAKTHROUGHS

Identification and characterization of dibasic proprotein convertases and discovery of the convertase-specific chaperone 7B2. Application: production of therapeutic products by animal cells.

MAJOR COOPERATIVE LINKS

In particular among the Universities of Cambridge, Leuven and Nijmegen, and the CLB. Joint experiments and exchange of information (workshops) and materials; recombinant proteins, polyclonal and monoclonal anti-convertase antibodies.

PUBLICATIONS

Joint publications

Van Duijnhoven, J.L.P., Kranenborg, M.G.C., Timmer, E.D.J., Groeneveld, A., Van den Ouweland, A.M.W., Creemers, J.W.M., Roebroek, A.J.M. and Van de Ven, W.J.M. Development and characterization of a panel of monoclonal antibodies against the el subtilisin-like proprotein processing enzyme furin. Hybridoma, **11**, 71-86, 1992.

Creemers, J.W.M., Roebroek, A.J.M., Van den Ouweland, A.W.M., Van Duijnhoven, J.L.P. and Van de Ven, W.J.M. Cloning and functional expression of a 4.3 kbp mouse FUR cDNA: evidence for differential transcription. Mol. Biol. (Life Sci. Adv.), **11**, 127-138, 1992.

Roebroek, A.J.M., Creemers, J.W.M., Pauli, I.G.L., Kurzik-Dumke, U., Rentrop, M., Gateff, E.A.F., Leunissen, J.A.M. and Van de Ven, W.J.M. Cloning and functional expression of Dfurin2, a subtilisin-like proprotein processing enzyme of Drosophila melanogaster with multiple repeats of a cysteine motif. J. Biol. Chem., **267**, 17208-17215, 1992.

Pelicci, G., Lanfrancone, L., Grignani, F., McGlade, J., Cavallo, F., Forni, G., Nicoletti, I., Grignani, F., Pawson, T. and Pelicci, P.G. A novel transforming protein with a SH2 domain acts as a competence factor for the cell cycle. Cell **70**, 93-104, 1992.

McGlade, J., Cheng, A., Pelicci, G., Pelicci, P.G. and Pawson, T. SHC proteins are phosphorylated and regulated by the v-src and v-fps protein tyrosine kinases. Proc. Natl. Acad. Sci. USA, **89**, 8869-8873, 1992.

Bailyes, E.M., Shennan, K.I.J., Smeekens, S.P., Steiner, D.F., Hutton, J.C. and Docherty, K.D. Characterization of the catalytic activity of the mammalian Kex2 homologue PC3 on proinsulin. Biochem. J. **285**, 391-394, 1992.

Rozakis-Adcock, M., McGlade, J., Mbamalu, G., Pelicci, G., Daly, R., Li, W., Batzer, A., Thomas, S., Brugge, J., Pelicci, P.G., Schlessinger, J. and Pawson, T. Association of the SHC and GRB2/SEM5 SH-2 containing proteins is implicated in activation of the RAS pathway by tyrosine kinase. Nature **360**, 689-692, 1993.

Rhodes, C.J., Thorne, B.A., Lincoln, B., Thomas, G. and Hutton, J.C. Specific processing of proopiomelanocortin by the insulin secretory granule proinsulin-converting type 2 endopeptidase. J. Biol. Chem., **268**, 4267-4275, 1993.

Creemers, J.W.M., Siezen, R.J., Roebroek, A.J.M., Ayoubi, T.A.Y., Huylebroeck, D., G. and Van de Ven, W.J.M. Modulation of furin-mediated proprotein processing activity by site-directed mutagenesis. J. Biol. Chem. **268**, 21826-21834, 1993.

Creemers, J.W.M., Groot Kormelink, P.J., Roebroek, A.J.M., Nakayama, K. and Van de Ven, W.J.M. Proprotein processing activity and cleavage site selectivity of the KEX2-like endoprotease PACE4. FEBS Lett. **336**, 65-69, 1993.

Segatto, O., Pelicci, G., Giuli, S., Digiesi, G., Di Fiore, P.P., McGlade, J., Pawson, T. and Pelicci, P.G. Shc products are substrates of erbB-2 kinase. Oncogene **8**, 2105-2112, 1993.

Dilworth, S., Brewster, C., Jones, M., Lanfrancone, L., Pelicci, G. and Pelicci, P.G. Transformation by polyoma virus middle T-antigen requires the binding and tyrosine phosphorylation of Shc. Nature, **367**, 87-89, 1994.

Huebner, K., Kastury, K., Druck, T., Salcini, A.E., Lanfrancone, L., Pelicci, G., Lowenstein, E., Li, W., Park, S-H., Cannizzaro, L., Pelicci, P.G. and Schlessinger, J. Chromosome locations of genes encoding human signal transduction pathway 'adaptor' proteins. Genomics, (in press), 1994.

Siezen, R.J., Creemers, J.W.M. and Van de Ven, W.J.M. Subtilisin-like proprotein convertases: Homology modelling of the catalytic domain of human furin. **222**, 255-266, 1994.

Individual publications

Creemers, J.W.M., Roebroek, A.J.M. and Van de Ven, W.J.M. Expression in human lung tumour cells of the proprotein processing enzyme PC1/PC3: Cloning and primary sequence of a 5 kb cDNA. FEBS Lett., **300**, 82-88, 1992.

Van de Ven, W.J.M., Creemers, J.W.M. and Roebroek, A.J.M. Furin, the prototype mammalian subtilisin-like proprotein processing enzyme. Enzyme, **45**, 257-270, 1992.

Bailyes, E.M. and Hutton, J.C. Kinetic analysis of the type I proinsulin endopeptidase using a monoclonal antibody-based immunoabsorbent assay. Biochem. J. **286**, 223-229, 1992.

Bennett, D.L., Bailyes, E.M., Nielsen, E. and Hutton, J.C. The identification of the type 2 proinsulin processing endopeptidase as PC2, a member of the eukaryote subtilisin family. J. Biol. Chem. **267**, 15229-15236, 1992.

Bailyes, E.M., Bennet, D.L. and Hutton, J.C. Proprotein processing endopeptidases of the insulin secretory granule. Enzyme **45**, 301-313, 1992.

Braks, J.A.M., Guldermond, K.C.W., Van Riel, M.C.H.M., Coenen, A.J.M. and Martens, G.J.M. Structure and expression of *Xenopus* prohormone convertase PC2. FEBS Lett. **305**, 45-50, 1992.

Powell, J.T., Klaasse Bos, J.M., van Mourik, J.A. The uptake and expression of the factor VIII and reporter genes by vascular cells. FEBS Lett. **303**, 173-177, 1992.

Powell, J.T., van Zonneveld, A.J., van Mourik, J.A. Gene transfer into specific vascualr cells. Eur J. Vasc. Surg. **6**, 130-134, 1992.

Guest, P.C., Bennett, D.L., Bailyes, E.M., Arden, S.D. and Hutton, J.C. The post-translational processing and intracellular sorting of the proinsulin type 2 endopeptidase PC2 in pancreatic islets. J. Biol. Chem., **267**, 22401-22406, 1992.

Van de Ven, W.J.M., Van Duijnhoven, J.L.P. and Roebroek, A.J.M. Structure and function of mammalian proprotein processing enzymes of the subtilisin class of serine proteases. Crit. Rev. Oncogenesis, **4**, 115-136, 1993.

Roebroek, A.J.M., Creemers, J.W.M., Pauli, I.G.L., Bogaert, T. and Van de Ven, W.J.M. Generation of structural and functional diversity in furin-like proteins in Drosophila melanogaster by alternative splicing of the DFUR1 gene. EMBO J. **12**, 1853-1870, 1993.

Mertens, K., Donath, M.J.S.H., van Leen, R.W., de Keyser-Nellen, M.J.M., Verbeet, M.Ph., Klaasse Bos, J.M., Leyte, A., van Mourik, J.A. Biological activity of recombinants variants lacking the central B-domain and the heavy-chain sequence Lys713-Arg740: discordant *in vitro* and *in vivo* activity. Brit. J. Haematol. **85**, 133-142, 1993.

Voorberg, J., Fontijn, R., Calafat, J., Janssen, H., van Mourik, J.A., Pannekoek, H. Biogenesis of von Willebrand Factor-containing organelles in heterologous transfected CV-1 cells. EMBO J. **12**, 749-758, 1993.

Lenting, P.J., Donath, M.J.S.H., van Mourik, J.A., Mertens, K. Identification of a binding site for blood coagulation Factor IXa on the light of human Factor VIII. J. Biol. Chem. **269**, 7150-7155, 1994.

Ayoubi, T.A.Y., Creemers, J.W.M., Roebroek, A.J.M. and Van de Ven, W.J.M. Expression of the dibasic proprotein processing enzyme furin is directed by multiple promoters. J. Biol. Chem., **269**, 9298-9303, 1994.

Arden, S.D., Rutherford, N.G., Guest, P.C., Curry, W.J., Bailyes, E.M. and Hutton, J.C. The post-translational processing of chromogranin A in the pancretic islet. Involvement of the eukaryote subtilisin PC2. Biochem. J., **298**, 521-528, 1994.

Bennett, D.L. and Hutton, J.C. Proinsulin-processing endopeptidases in 'Molecular Biology of Diabetes' Draznin, B. & Leroith, D. (eds) Humana Press Inc, New York, (in press).

Hutton, J.C. Insulin-secretory granule biogenesis and the proinsulin-processing endopeptidases. Diabetologia, (in press), 1994.

Martens, G.J.M., Braks, J.A.M., Eib, D.W., Zhou, Y. and Lindberg, I. The neuroendocrine polypeptide 7B2 is an endogenous inhibitor of the prohormone convertase PC2. Proc. Natl. Acad. Sci. USA. **91**, 5784-5787, 1994.

Braks, J.A.M. and Martens, G.J.M. 7B2 is a neuroendocrine chapaerone that transiently interacts with prohormone convertase PC2 in the secretory pathway. Cell **78**, 263-273, 1994.

Patent applications

'The 7B2 protein and recombinant DNA and cDNA, mRNA, and antibodies for 7B2 and protocol for the detection of 7B2'27; patent application 8702590, Stichting Katholieke Universiteit Nijmegen; The Netherlands, 30 October 1987; Europe, Japan, USA (application 07/263,885), **28** October 1988; inventor: G.J.M. Martens.

'The 7B2 protein as a chaperonin — Production of bioactive peptides with recombinant cells'; patent application 91.02009, Stichting Katholieke Universiteit Nijmegen; The Netherlands, 29 November 1991; inventor: G.J.M. Martens. The related patent 'Production of proteins' has been filed for Europe, USA and Japan by Ciba-Geigy Ltd, Basel, Switzerland (30 November 1992).
'7B2 as enzyme inhibitor'; patent application 94.00032, Stichting Katholieke Universiteit Nijmegen; The Netherlands, 7 January 1994; inventor: G.J.M. Martens.

Development of an integrated vector system for the expression of immunoglobulins in different cellular compartments of mammalian cells (BIOT CT-920306)

COORDINATOR:
A. BRADBURY, Società Italiana per la Ricerca Scientifica, Roma, I
PARTICIPANTS:
A. CATTANEO, International School for Advanced Studies, Trieste, I
H. HOOGENBOOM, Cambridge Antibody Technology, Melbourn, GB

BACKGROUND INFORMATION

This project is aimed at bringing together the two fields of phage antibodies and the ectopic expression of antibodies in animal systems in a way which will provide an innovative and general approach to the study and manipulation of animal cells for scientific or biotechnological purposes.

Genetic manipulation of antibody molecules has progressed considerably in recent years. FAb fragments, Fv fragments, isolated VH domains and single chain Fv fragments (ScFv) have all been expressed and secreted into the periplasm of Escherichia Coli. ScFvs (as well as many other proteins) have also been expressed on the head of bacteriophage fd by incorporating the genome for the ScFv at the 5' end of the fd gene 3. This couples the expression of the antibody specificity with the genetic information encoding that specificity. The creation of large libraries of phage antibodies, with each phage expressing a different antibody, has allowed the selection of antibodies which bind to a particular antigen. One great advantage of this system is that the cloning of the antibody gene is simultaneous with the isolation of the antibody with the desired binding activity.

Expression of recombinant antibodies in a variety of mammalian cells as secreted or intracellular proteins targeted to different cellular compartments has been shown to occurr efficiently. This now opens the possiblity of intracellular or intercellular 'immunisation' by recombinant means as a general strategy to inhibit cellular functions or to introduce new phenotypes in animal cells.

In this project we have created an integrated system which will allow cloning of any antibody variable region of interest, either from an already established hybridoma, or by selection from an appropriate library, into one of a panel of vectors which will provide the necessary flanking sequences (promoter, cellular localisation information, extra effector functions) for the particular purpose desired. This will allow a researcher to answer a particluar problem by selecting an appropriate form of antibody to be expressed in the desired cellular compartment of a selected cell.

OBJECTIVES AND PRIMARY APPROACHES

The final objectives of this EC project are the following:

(1). The development of the phage (and phagemid) vectors for the creation of the antibody libraries.

(2) The creation of phage antibody libraries of diversity high enough (greater than 10^6) to select useful antibodies.

(3) The study of the stability of different forms of antibody binding domains (single chain, FAb, single chain attached to constant regions, complete antibody) within the cellular compartments of mammalian cells. This will be done with antibodies presently under study (Y13/259, anti ras p21; αDII, anti nerve growth factor; phox, antiphenoxazolone.

(4) The testing of targeting signals as applied to antibody domains.

(5) The creation of 'targeting signal cassettes', which by the addition of appropriate sites will allow straightforward cloning into the integrated vectors.

(6) The development of mammalian plasmid expression vectors containing restriction sites compatible with those found within the phage vectors (in 1), containing the targeting cassette to direct the antibody to the appropriate celluar compartment (4), and expressing the form of antibody found to be most stable within the cellular compartment under study (3).

(7) The development of similar vectors as in 6, but (retro)viral in origin.

(8) The testing of the ability of the general vectors to direct correct expression of antibody domains.

(9) Functional analysis of cell lines expressing antibody domains directed against the selected intracellular or extracellular antigens presently under study, and others which may be selected from a phage library. General vectors will be used for this goal when they become available.

RESULTS AND DISCUSSION

Phage and phagemid vectors have been engineered and used to create phage antibody libraries with diversities greater than 10^6. Preliminary studies on the stability of different antibody forms have been performed which show that, in general, whole antibodies are probably better for the expression of antibodies in extracellular compartments (extracellular, endoplasmic reticulum, cell membrane) while ScFvs are better intracellularly (cytoplasmic, nuclear, mitochondrial), probably because of their smaller size. Antibodies (or derivatives) have been targeted to the following cellular compartments: secretory, endoplasmic reticulum, cytoplasmic, nuclear and mitochondrial. A backbone vector which contains a polylinker, eukaryotic selectable markers (neomycin or GPT), prokaryote selectable marker (β lactamase), colE1 origin of replication, fd origin of replication and SV40 origin of replication all flanked by restriction sites which allow their easy removal and replacement has been created and shown to be functional. A heavy chain cassette, containing leader, V region polylinker (compatible with sites in the phage antibody vectors), and human IgG1 gene has also been produced. In both of these vectors many restriction sites were removed by *in vitro* mutagenesis to ensure that useful sites were unique. Both heavy and light chain cassettes were shown to be functional when inserted into the backbone (VHexpress & VLexpress). Derivatives of these plasmids which will allow efficient expression of complete immunoglobulins from a single plasmid (IGexpress) will also be produced by the end of this project. Targeting signals as cassettes have been produced, incorporated into plasmids and are in the process of being tested.

In summary, an integrated vector system will be produced and tested by the completion of the project. This will allow easy expression of antibody V regions, either as complete antibodies or as derivatives, in different cellular compartments.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The expression of antibodies (or domains) in animal cells will confer new phenotypes on those cells either in tissue culture or in the whole animal (transgenic animals). These phenotypes will express themselves as either loss or gain of functions which will be of research interest, but will also have biotechnological applications. For example, protection of livestock from infectious disease by expression of specific antibodies against pathogens or viruses in their serum (a transgenic passive form of immunisation), in mothers milk (using antibody forms resistant to proteolysis) or in crops used for feed (phytoantibodies) would be relatively easy using our proposed system. Furthermore, it may be possible to create models for animal or human disease by the tissue-specific expression of suitable antibodies. Many other applications of this strategy can be envisaged. A vector system for the expression of phage antibodies has been marketed. The vector system developed within the framework of this project is compatible with this and will be probably be marketed as an add on module.

MAJOR COOPERATIVE LINKS

In the eighteen months since we started this project have met six times to discuss the exact details of the vector system and have communicated numerous times by telephone and fas. There have been three meetings with the other members of the T project (once in Ireland, once in Portugal and once in Germany), as a result of which, we have received DNA samples from two groups (Bujard and Grummt & Hauser) which we are using in the vector system.

PUBLICATIONS

Joint publications

Bradbury, A., Ruberti, F., Werge, T., Amati, V., Di Luzio, A., Gonfloni, S., Hoogenboom, H., Biocca, S., & Cattaneo, A. (1994). The cloning of hybridoma V regions for their ectopic expression in intracellular and intercellular immunisation. In C. Borrebaeck (Ed.), *Antibody Engineering II: a Practical Approach*, New York: IRL Press.

Ruberti, F., Bradbury, A., & Cattaneo, A. (1993). Cloning and expression of an anti-NGF antibody for studies using the neuroantibody approach. *Cell. Mol. Neurobiol.*, **13**, 559-568.

Ruberti, F., Cattaneo, A., & Bradbury, A. (1994). The use of the RACE method to clone hybridoma cDNA when V region primers fail. *J. Imm Methods*, **173**, 33-39.

Construction of permanently transfected cells expressing steroid hormone receptors (BIOT CT-920308)

COORDINATOR:
A. MAGGI, University of Milan, Milan, I
PARTICIPANTS:
H. BUJARD, University of Heidelberg, Heidelberg, D
D. DI LORENZO, University of Brescia, Brescia, I

BACKGROUND INFORMATION

The study of the mechanism of action of steroid receptors has recently received wide interest because these receptors are hormonally-inducible transcription factors and therefore represent an important tool to elucidate the bases of the regulation of gene activity.

Our interest in the field was restricted to one major point that is the elucidation of the mechanisms involved in the tissue specificity of action of these class of receptors. It was therefore proposed to generate a series of cell lines expressing steroid receptors. In view of the possible toxic effect of these receptors, their synthesis had to be regulatable. Moreover, the control of the intracellular concentration of the receptors was also of interest since it would allow the analysis protein-dosage effects.

Because of the well described limits of eucaryotic regulatable promoters, for the present study it was proposed to utilize also a heterologous regulatory system recently developed in H. Bujard's laboratory. This system, in fact, has been shown to be very efficient. In brief, the system consists of a transcriptional transactivator which is a fusion protein between the repressor of the Tn10-derived tetracycline resistance operon of *E. coli* and the activation domain of Herpes Simplex Virus protein 16. The second element of the system is a minimal promoter derived from the cytomegalovirus promoter IE, fused downstream of multiple tet operator sites. In many cell lines, the activity of this promoter is ($P_{CMV^{*}-1}$) entirely dependent on the binding of tetracycline responsive transactivator (tTA). The gene coding for the protein to be expressed therefore has to be inserted downstream of $P_{CMV^{*}-1}$. In order to generate cell lines synthesizing steroid receptors in a regulated fashion we had to generate first cell lines which stably produced the transactivator. A second stable transfection was then performed with DNA encoding the steroid receptor under the control of $P_{CMV^{*}-1}$.

OBJECTIVES AND PRIMARY APPROACHES

The objective of this research was to utilize a newly developed regulatory system for the tight control of the expression of steroid receptor genes.

RESULTS AND DISCUSSION

1. Cell lines generated

a. Lines expressing the transactivator

Three cell lines of different tissue origin were utilized for the transfection: rat hepatoma (HTC); human breast cancer (T47D); human neuroblastoma (SK-N-BE).

Several clones for each cell line were selected in conditioned medium, then expanded and tested for the tTA activity. This analysis was performed by transient transfection, in each clone, of a reporter represented by the luciferase gene under the control of P_{CMV*-1}. The luciferase activity was analyzed in the presence or absence of tetracycline.

These experiments allowed to conclude that the P_{cmv*-1} driven production of luciferase was similarly regulated in all the cell lines generated.

b. Cell lines expressing steroid receptors

Utilizing an expression vector with a metallothionein promoter two cell lines were generated:
human neuroblastoma (SK-N-BE) expressing the progesterone receptor; human neuroblastoma (SK-N-BE) expressing the estrogen receptor

2. Analysis of the activity of the stably transfected genes

a. Analysis of the specificity of action of the transactivator

A series of studies were carried out in order to assess the specificity of the activity of the transactivator. The clones with the highest transactivator activity where then analyzed in order to determine whether the presence of the fusion protein could interfere with the transcriptional activity of endogenous genes as well as genes transiently transfected.

The proteins initially studied were C/EBPb, glucocorticoid receptor and SP1.

To our surprise we found that in the presence of tetracycline the intracellular levels of the gene products were highly decreased (with respect to the parental cell line) while the removal of tetracycline resulted in a generalized increase of the levels of these proteins.
We then tested the effect of tetracycline presence and removal on the expression of c-fos and Insulin-like Growth Factor I Receptor. The levels of these two proteins were unaffected by the presence/absence of tetracycline.

It was concluded that the presence of the transactivator may interfere with the host cell transcriptional activity, but the effect is not generalized.

b. Analysis of the cellular effects of the presence of the estrogen and progesterone receptor

The neuroblastoma cells stably transfected with the estrogen and progesterone receptor (ER and PR, respectively) were initially tested in order to assess the range of concentrations of expression of the two receptors and their viability. Clones containing a physiological concentration of the estrogen (about 300-400 fmoles/mg protein) and progesterone (200 fmoles/mg protein) receptors were expanded and further tested.

A growth and morphology analysis of the cells expressing the two receptors demonstrated that the activation of the progesterone rece ptor did not change these parameters, while the activation of ER determined growth arrest and morphological differentiation of the cells. Further analysis of the ER-expressing cells (named SK-ER3) demonstrated that in this cell line the activation of the estrogen receptor causes a profound morphological and biochemical differentiation towards a phenotype highly reminiscent of the mature neuron.

MAJOR SCIENTIFIC BREAKTHROUGHS

a. Some of the cell lines generated can be utilized for the screening of compounds acting on estrogen or progesterone receptors.

b. A heterologous transcription control system for eucaryotic cells was utilized to generate a series of cell lines suitable for the regulated synthesis of steroid receptors.

c. The generation of a human neuroblastoma cell line expressing the estrogen receptor gene allows the analysis of the effect of estrogen in cells of nervous origin. This represents an important achievement since the study of the effect of estrogens on nervous cells has been traditionally hampered by the lack of an appropriate model system. Furthermore, our studies demonstrate that the activation of the estrogen receptor leads to the differentiation of cells of nervous origin towards the phenotype of mature neurons. Therefore the cell system generated besides representing a tool for the analysis of estrogen action in neural-derived cells, opens new perspectives for the possible role of this sex steroid in the development of nervous cells.

MAJOR COOPERATIVE LINKS

Exchange visits of the project participants. Exchange of reagents and know-how among the participating laboratories: the tet regulatory system was developed in H. Bujard's laboratory. Five vectors which allowed to establish this regulatory system in the laboratories of A. Maggi and Di Lorenzo were supplied from Heidelberg. In addition, a vector containing the coding sequence of the human estrogen receptor under the control of P_{CMV^*-1} was sent from the Heidelberg Lab to Di Lorenzo's group. The regulatory system was established in a number of cell lines and studied with regard to the regulated expression of seroid receptor genes in a number of cell lines. These studied were carried out in collaboration between A. Maggi's and D. Di Lorenzo's laboratories.

PUBLICATIONS

Joint publications

Magalini A., Ferrari F., Savoldi G., Ingrassia R., Pollio G., Maggi A. and Di Lorenzo D. Specificity of action of Herpes virus VP16-tet transactivator in mammalian cell cultures. (in preparation)

Individual publications

Ma Z.Q., Bondiolotti G.P., Olasmaa M., Violani E., Patrone C., Picotti G.B. and Maggi A. Estrogen Modulation of catecholamine synthesizing monoamine oxidase A activity in the human neuroblastoma cell line SK-ER3. *J. Steroid Biochem. Mol. Biol.* **47**: 207-211, (1993)

Santagati S., Bettini E., Asdente M., Muramatsu M. and Maggi A. Theoretical considerations for the application of competitive PCR to the quantitation of a low abundance mRNA: estrogen receptor. *Biochemical Pharmacology,* **46**: 1797-1803, (1993)

Ma Z. Q., Spreafico E., Pollio G., Santagati S., Cattaneo E. and Maggi A. Activated estrogen receptor mediates growth arrest and differentiation of a neuroblastoma cell line. *Proc. Natl. Acad. Sci. (USA)* **90**: 3740-3744, (1993)

Santagati S., Melcangi R.C., Celotti F., Martini L. and Maggi A. Estrogen receptor is expressed in different types of glial cells in culture *J. Neurochemistry,* (in press)

Ma Z.Q., Santagati S., Patrone C., Pollio G., Vegeto E. and Maggi A. Insulin-like growth factors activate estrogen receptor to control the growth and differentiation of the human neuroblastoma cell line SK-ER3. *Molecular Endocrinology,* **7**: 910-918, (1994)

Di Lorenzo D., Giannì M., Savoldi G., Ferrari F., Albertini A. and Garattini E. Progesterone induced expression of alkaline phosphatase is associated with a secretory phenotype in T47D breast cancer cells. Biochem. *Biophys. Res. Commun.,* **192**: 1066-1072, (1993)

Ingrassia R., Savoldi G., Garaffini A., Tironi M., Poiesi C., Williams P., Albertini A. and Di Lorenzo D. Analysis of a novel transcription complex that cooperates with the glucocorticoid receptor for the hormone regulation of the alpha-1-acid glycoprotein (AGP) gene. *DNA and Cell Biology*, vol. 13 n° 6, (1994).
Savoldi G., Ferrari F., Ruggeri G., Sobek L., Albertini A. and Di Lorenzo D. Progesterone agonists and antagonists induce down- and upregulation of estrogen receptors and estrogen inducible genes in human breast cancer cell lines. *The International Journal of Biological Markers*, (in press)
Gossen M. and Bujard H. Tight control of gene expression in mammalian cells by tetracycline responsive promoters. *Proc. Natl. Acad. Sci. (USA)*, **89**: 5547-5551, (1992)
Gossen M. and Bujard H. Anhydrotetracyclin, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells. *Nucleic Acid Res.*, **21**: 4411-4412, (1993)
Gossen M., Bonin A. and Bujard H. Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. *TIBS*, **18**: 471-475, (1993)
Weinmann P., Gossen M., Hillen W., Bujard H. and Gatz C. A chimeric transactivator allows tetracycline-responsive gene expression in whole plants. *Plant J.*, **5**: 559-569, (1994)
Resnitzky D., Gossen M., Bujard H. and Reed S.I. Acceleration of the G1/S phase transition by expression of cyclins D1 and E using an inducible system. *Mol. Cell. Biol.* **14**: 1669-1679, (1994)
Furth P.A., Onge L.St., Böger H., Gruss P., Gossen M., Kistner A., Bujard H. and Hennighausen L. Temporal control of gene expression in transgenic mice by a tetracycline responsive promoter. *Proc. Natl. Acad. Sci. (USA)*, **91**: 9302-9306, (1994)
Früh K., Gossen M., Bujard H., Peterson P.A. and Yang Y. Displacement of housekeeping proteasome subunits by MHC-encoded LMPs: a new mechanism for modulating the multicatalytic proteinase complex. *EMBO J.*, (in press)

Targeted inducible amplified homologous expression system for quality products from animal cells in culture (BIOT CT-920307)

COORDINATOR:

A: F. GANNON, Nat. Diagnostics Centre, Univ. College Galway, Galway, IRL

PARTICIPANTS:

B: M. BIANCHI, Univ. of Milano, Milano, I

C: M. CLYNES, Nat. Cell & Tissue Culture Centre, Dublin City Univ. Glasnevin, Dublin, IRL

D: J. LEHMANN, University of Bielefeld, Cell Culture Technology, Bielefeld, D

E: J. FREY, Univ. Bielefeld, Bielefeld, D

F: E. HOUNSELL, Univ. College, London, GB

G: C. ROSSI, IZS, Inst. Zooprofilattif, Brescia, I

H: K. BERGEMANN, K. Thomae GmbH, Biotechn. Produktion, Biberach an der Riss, D

BACKGROUND INFORMATION

The aim of this project was to integrate gene expression (A, B, E, G, H), post-translational modification (F) and cell culture studies (C, D, H).

OBJECTIVES AND PRIMARY APPROACHES

Groups B and G were active in attempting to obtain targeted and amplifiable expression systems. The CAD gene which was selected for this is amplifiable by the innocuous compound PALA. Any internally targeted exogenous gene should be co-amplified thus allowing high levels of expression. Group A was looking for an inducible homologous expression system from CHO. The experimental approach used involved the preparation of a cDNA library from CHO cells grown in the absence of serum followed by differential hybridisation of this library using total cDNA from cells grown either in the presence or absence of serum (provided by groups C and H) as probes. The promoter sequences were isolated by group A from a generated CHO genomic library (genomic DNA provided by group H). Post-translational studies were carried out by group F on a model antibody. The expression system for this model protein was set up in the laboratories E and H and the antibody was produced in the laboratories D and H. Groups C, D and H studied the impact of serum depletion (C, H) and amino acid deprivation (D) on gene expression (in cooperation with groups A and E) and cell growth and the conditions which are necessary to successfully transfer cells from full serum to serum-free media (C).

RESULTS AND DISCUSSION

The CAD gene was isolated and sequenced from CHO. A construct was prepared in which the neomycin resistance gene was carried within the intron of this amplifiable gene. This construct was used to introduce the neo marker by homologous recombination into the cellular CAD gene locus. A number of neoresistant cells were obtained following transfection. However, after the analysis

of over 1000 potentially positive cell clones none were found to have the gene integrated in the correct locus.

Following differential hybridisation several highly expressed and induced cDNA clones were isolated. Analysis of the sequences showed that many of these were derived from mitochondria, others came from ribosomal proteins, others from ubiquitin related genes and some corresponded to previously undescribed genes. The inducible genes were all encoded in the mitochondrion. It was shown that this was due to an amplification of mitochondria per se rather than to an induction at the level of transcription of mitochondria. A further increase in mitochondrial transcripts was observed using RNA obtained from serum-free cultivated cells which were shifted from a complete amino acid medium to one which was deprived from an non essential amino acid . The promoters of the chromosomal genes have been isolated and by the end of the time period of the project will be available for tests on their efficiency.

The expression of the model antibody used for these studies gave rise to a lower level of sialylation when cells were grown in culture compared to commercial polyclonal serum immunoglobulin. These model antibody preparations would therefore be expected to have a shorter half-life in circulation. Cell culture studies are being undertaken to obtain conditions in which this pattern is altered.
The studies on cell cultivation conditions gave rise to new information on the cell line which is most frequently used in the biotechnology industry. CHO cells were adapted successfully to growth in optimised serum-free media. Their tolerance to a variety of environmental conditions such as agitation rates and levels of the waste products lactate and ammonia has been investigated. Analysis on the amino acid requirements of CHO showed that the cells could compensate for the depletion of the non essential amino acids serine, asparagine and glycine whereas cysteine and proline depletion led to growth arrest and tyrosine deficiency to cell death.

MAJOR SCIENTIFIC BREAKTHROUGHS AND INDUSTRIAL APPLICATIONS

The presence in this group of an industry has ensured the transfer to the biotechnology industrial sector of the data whch has been obtained. A direct consequence of this is a patent which is currently being prepared to cover the newly described CHO promoter sequences. The data were also widely disseminated at the ACTIP meetings and at the T-project meetings which were held in Galway, Lisbon and Bielefeld. Of particular interest in this context were the culture conditions which could be fused for the growth of CHO cells.

MAJOR COOPERATIVE LINKS

The exchange of material which was necessary to achieve the work described above is obvious from the description of this work. In addition there were exchanges of personnel between laboratories (C and D, A and E, A and H). The exchanges of material were between all of the different participants with a particularly strong input in the transferal of RNA from groups C, D and E to group A and the transfer of the CHO cell line from group H to groups A, B, C, D, E and G. The cell lines which contained the model expression system were prepared by groups E and H and worked upon by group F.

PUBLICATIONS

Joint publications

M. J. Davies, K. Bergemann, W. MacDowell and E.F. Hounsell: The glycosylation pattern of a humanised. IgG1 antibody (D1.3) expressed in CHO cells (in preparation)

B. Enenkel, K. Bergemann and F. Gannon: Increased transcript level of mitochondrial genes in serum-free cultivated CHO cells due to increase in the number of mitochondria (in preparation)
T. Seewoester, F. Korioth, J. Frey and J. Lehmann: Non-essential amino acids — exemplified by serine and asparagine — as modulators of the transcription activity in chinese hamster ovary cells (in preparation)

Individual publications
L. Byrnes, B. Enenkel, F. Gannon and T. Smith: Preparation of DNA probes In Gene Probes — Vol. I: A Practical Approach, Eds. B.D. Hames and S.J. Higgins, IRL Press (in press)

T-PROJECT

'PLANT REGENERATION'

Factors regulating growth and differentiation of plant cells. A basis for the understanding of plant regeneration (BIOT CT-900206, 900160, 900158, 900177, 900179, 900178)

COORDINATOR:
J. GUERN, Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, F
PARTICIPANTS:
K. PALME, Max-Planck-Institut, Köln, D
M. HALL, Univ. College Wales, Aberysthwyth, GB
P. COSTANTINO, Univ. di Roma 'La Sapienza', Roma, I
S. de VRIES, Agricultural Univ. Wageningen, NL
P. PECHAN, Max-Planck-Institut, Biochem., München, D
A. van LAMMEREN, Agric. Univ. Wageningen, Wageningen, NL
A. MARTIN, CSIC, Cordoba, E
A. McQUEEN, Nickersons, Cambridge, GB
H. DONS, CPRO-DLO, Wageningen, NL
K.R. LIBBENGA, University of Leiden, Leiden, NL
D. KLÄMBT, University of Bonn, Bonn, D
P. ADUCCI, University of Rome 'Tor Vergata', Rome, I
P. van den ELZEN, Mogen Int. NV, Leiden, NL
H.A. VERHOEVEN, CPRO, Wageningen, NL
M. TERZI, Univ. Padova, Padova, I
K. ROBERTS, JII, Norwich, GB
J.D. MIKKELSEN, Maribo, Copenhagen, DK
S.C. FRY, Univ. Edinburgh., Edinburgh, GB
A. SPENA, Max-Planck-Institut, Köln, D
J. TEMPÉ, Institut Sciences Végétales CNRS, Gif-sur-Yvette, F
H. VAN ONCKELEN, Univ. Instelling Antwerpen, Antwerp, B
D. CHRIQUI, Cytol. Embriol. Végétales. Univ. P. & M. Curie, Paris, F
M. BOULAY, L. Vuitton Moet Hennessy Recherche, Paris, F
D. MARIOTTI, Istituto Radiobiol. Ecofisiol. Vegetali CNR, Monterotondo, I
M. VENIS, Horticulture Research Int., Kent, GB
H. BARBIER-BRYGOO, CNRS, Gif-sur Yvette, F
D. INZÉ, Rijksuniversiteit Gent, Gent, B
P.A. MILLNER, University of Leeds, Leeds, GB
F. OKKELS, Danisco/AS, Copenhagen, DK

BACKGROUND INFORMATION

The regeneration of plants from protoplasts, cells or plant organs is of utmost importance in the use of modern biotechnology for plant improvement. Model species in terms of easiness of regeneration have been extensively used and a significant number of plants of agricultural interest have been regenerated. Unfortunately, the procedures used are largely empirical and fail to be transferred to the so-called recalcitrant species or varieties. This means that the level of basic knowledge concerning the mechanisms that underlie the developmental steps involved in the regeneration processes is critically limiting. The most often used empirical strategy to regenerate plants is to manipulate the concentrations and balance of auxin and cytokinins, two growth factors known to trigger cell division and organogenesis. A deaper knowledge of the mechanisms of action of these two hormones as well as that of other growth factors known to influence cell differen-

tiation would bring strategic information in order to build rational approaches of regeneration. On another hand, a variety of plants exhibit the remarkable property to be able to generate embryos either from somatic cells or from microspores diverted form their normal pathway of pollen differentiation. These remarkable properties offer unique possibilities for studying the regulation of embryo development as well as promising biotechnological applications in terms of regeneration of newly created genotypes. Unfortunately, success in obtaining somatic embryos or embryos derived from microspores concerns a limited number of species and here again getting a better knowledge of the cellular and molecular basis of these developmental steps is critical.

OBJECTIVES AND PRIMARY APPROACHES

This T-Project has been built with the aim to give a strong impetus in Europe to the studies concerning the mechanisms of action of several signals regulating growth and differentiation of plant cells and to get through these studies a better understanding of the regulation of the initial steps of plant morphogenesis. The general strategy has been to associate in the same multidisciplinary group scientists working on the basic mechanisms of growth factor perception at the cellular level with those involved in the study of key morphogenetic events such as somatic embryogenesis, microspore embryogenesis and rhizogenesis, the ultimate target being to improve the basic knowledge necessary for rendering the regeneration of plants a more predictable process.

The task has been distributed between five groups of comparable size which have been involved in five contracts addressing the mechanisms of action of plant growth factors (BIOT CT-900158: *Perception, Interactions and Responses to Plant Growth Regulators*; BIOT CT-900178: *Molecular Analysis of Auxin-specific Signal Transduction in Plant Cell Communication*), the regulation of embryogenesis (BIOT CT-900160: *Regulation of the Inductive Phase of Microspore Embryogenesis*; BIOT CT-900177: *Molecular Analysis of Higher Plant Embryogenesis*) and the use of microbial genes to study plant morphogenesis (BIOT CT-900179: rol *Genes as Privileged Tools to Study Plant Morphogenesis).*

RESULTS AND DISCUSSION

Some of the major scientific achievements are highlighted here but a detailed description of the progress achieved in each programme of the five groups involved in the T-Project can be found in the final reports of the corresponding contracts.

A major effort, which is one of the originality of the european research in Plant Sciences, has been devoted to the study of proteins putatively involved in the perception of growth factors and related molecules. The purification and partial or complete sequencing of new auxin-binding proteins (ABPs), ethylene-binding protein (EBPs) and fusicoccin-binding proteins (FCBPs) have been achieved. These efforts and their results give to the european plant biologists quite a strong position in the domain. The most systematic description ever done of auxin-binding proteins has been undertaken with the aim to identify genes involved in auxin perception. A variety of auxin-binding proteins with different cellular locations (soluble, plasma membrane, plasmodesmata, peribacteroidal membranes, etc.) have been characterized. ZmER ABP1, the major auxin-binding protein from the endoplasmic reticulum of maize coleoptiles, is by far the best characterized protein able to bind auxin. Immunological probes directed against epitopes of this protein have been built and characterized, allowing to explore functions of ZmER ABP1

or related proteins. Synthetic peptides corresponding to the presumed auxin-binding site of ABPs have been prepared and the corresponding antibodies raised and characterized in terms of activity. Most remarkably, one of them, D 16, exhibits an auxin agonist activity and represents the first antibody with agonist activity ever evidenced in plants. The study of the protein ZmER ABP1 has now reached the state where it has been dissected in its peptide subcomponents, mapped for their biological roles. It has been discovered that the 14 aminoacid C-terminal peptide of ZmERABP1 inhibits inward K^+ channels and stimulates outward K^+ channels in *V. faba* guard cells, likely by triggering a cytoplasmic alkalinization. Another remarkable issue of these studies has been the discovery of a soluble 60kD protein from maize, able to bind auxins but hydrolyzing cytokinin conjugates, the hydrolysis being inhibited by IAA glucoside. The interactions between auxins and cytokinins which are of tremendous importance in the regeneration of plants from cell or organ cultures are very poorly known in terms of mechanisms; the results obtained here represent the first clear identification at the molecular level of one target where the two hormones interact.

For ethylene-binding and fusicoccin-binding proteins too, the T-group had a leading international impact. Despite very considerable difficulties resulting from the hydrophobic nature of the proteins involved and their extremely low abundance in plant tissue the receptors for both ethylene and fusicoccin were purified to homogeneity and partial sequences obtained. EBPs have been purified, separated into subcomponents against which antibodies have been prepared. Partial sequences have been obtained showing no homology with any known protein, gene cloning is in progress. Interestingly, evidence has been obtained suggesting that EBPs can be phosphorylated by an ethylene-dependent mechanism.

Purification of fusicoccin-binding proteins has been achieved, partial sequences obtained, specific antibodies produced. The fusicoccin receptor resembles the class of proteins known as 14-3-3 that act as regulators of protein kinases in many eukaryotes. Dual reconstituted systems with purified ATPase and partially purified FC receptors have been shown to be functionally responsive to FC. In this domain again, the group has had a strong leading position.

Another area of strong european leadership concerns the discovery that extracellular proteins are produced by plant cells, with specific patterns associated with cell differentiation. Really exciting results have been obtained concerning the role of some of these secreted proteins in the regulation of somatic embryo development in carrot. The characterization of extracellular proteins capable of rescuing the temperature sensitive mutant *ts11* has demonstrated the role of a specific extracellular 32 kD endochitinase. These results raise the general question of the targets of the extracellular proteins controlling cell differentiation. In this respect, the recent discovery that Nod-factors mimic the effect of the 32 kD endochitinase in rescuing the *ts11* mutant is really exciting in opening the possibility that regulators more or less similar to bacterial Nod factors might be produced by plant cells. Large scale purification of several extracellular chitinases secreted by carrot cell suspensions has been performed, allowing their biochemical characterization, the production of antibodies, the study of their activity and the cloning of the corresponding cDNAs.

The role of the extracellular matrix in cell differentiation and plant development has progressively emerged as a topic of major interest during the past few years. Research has developed along two lines. First, a collection of monoclonal antibodies has been raised against cell wall or cell surface epitopes and used as a tool to reveal cell and tissue differentiation steps. More specifically, monoclonal

antibodies raised against cell wall arabinogalactan-proteins (AGPs) have been used to monitor cell differentiation in *Arabidopsis* roots, as expressed by changes in the properties of the cell surface. Developmentally regulated patterns of cell surface expression of AGP epitopes have been observed. Even more interesting is the fact that cell lineages can be traced from the presence and absence of specific epitopes. AGP families thus provide molecular markers present at the surface of plant cells for studying the formation of the patterns of tissues in the plant. Even more interesting is the hypothesis raised by these results that AGPs could have a role in cell interactions or cellular signalling. The second line of research has focused on important enzymes of the cell wall, that could be involved in the regulation of cell growth and in the generation of oligosaccharidic signals generated within the wall. New active oligosaccharides of the xyloglucan type have been characterized. A new enzyme activity has been discovered in the plant cell wall, catalyzing endotransglycosylation between pairs of long xyloglucan-related molecules. This enzyme is likely involved in wall loosening and consequently in cell enlargement.

Another area where the group has brought a quite significant progress concerns the regulation of cell division in plants. Several *Arabidopsis* genes with high homology with cell cycle regulation genes have been cloned. Two *cdc2* genes have been characterized from their capacity to complement a conditional mutant of yeast affected in the *p34cdc2* kinase. Five different cyclin genes have also been characterized in *Arabidopsis*. The availability of the new tools corresponding to *cdc2* and cyclin genes has been the basis of new studies on the regulation of the cell cycle in plants. Specific expression patterns of these genes have been recognized. For the first time the influence of auxins and cytokinins on their expression has been studied, in relation with the foci of cell division. Interestingly, mutant *cdc2* genes overexpressed in *Arabidopsis* and tobacco strongly modify cell division. Strong decreases in the number of cells per organs with or without increases in cell size have been observed with mutant D-N *cdc2* gene in tobacco, and, surprisingly, this occurs without alterations in the general pattern of organ forming. These results clearly demonstrate that cell division and morphogenesis can be largely independent and provide new perspectives to alter plant development. This important achievement of the T-Project has to be considered in relation with the history of this programme. In 1989, attempts to launch a special effort concerning the cell cycle in plant cells were unsuccessful. Hopefully things have maturated and specific efforts can now be built with the objective to start a new description of plant morphogenesis in terms of regulation of the expression of genes involved in the regulation of the cell cycle. This is a good example of a domain where basic research rendering available new basic tools can bring essential information on the initial steps of regeneration processes.

The mechanisms by which *rol* genes of the T-DNA of *Agrobacterium rhizogenes* modify plant morphogenesis has been a subject of rather intense investigations, with stimulating new results. The *rolB* promoter has been dissected and the control of the gene expression studied in details. Five domains have been recognized in the promoter as responsible for directing the expression of the gene in different meristematic cells. The gene constructs corresponding to modified *rolB* promoters associated with the GUS sequence show that those promoters able to direct the expression in meristematic cells are auxin-inducible whereas those which determine expression in non-meristematic cells are not. A quite new and important finding has been to show that *rolB* appears as a 'meristem-inducing gene' rather than a root-inducing gene as it was classically admitted before. The type of meristem induced depends on the organ/tissue from which the meristems are induced; flower meristems can be induced under the influence of *rolB* expression, provided floral

stalks are used. The function proposed for the *rolB* gene, that of a glucosidase involved in the generation of free auxin from glucose ester conjugates, has been a matter of debate inside and outside the group and this hypothesis is now dismissed. This debate had an important 'contaminating' effect, stimulating research in laboratories outside the project and initiating a renewal of the interest for hormone metabolism in its relation with plant development. The possibility to exploit *rol* gene-induced morphogenetic alterations in plants of economical interest has also been explored by several groups in the T-Project. Kiwi, chestnuts, and alfalfa expressing *rolB* and rose plants expressing *rolC* have been studied. In alfalfa, increases in foliage and rooting systems have been observed whereas in roses, interesting new architectures have been obtained. Strickingly, the group has also shown that the Agrobacteria have still new things to reveal in terms of genes and gene products able to affect plant morphogenesis. This is examplified by the discovery of ORF13, a newly recognized gene from *A. rhizogenes*, which induces new phenotypic alterations in transgenic tobacco.

The project relative to the induction of microspore embryogenesis from immature pollen used an agronomically important crop, *Brassica napus*, as a model. In this material, a simple temperature treatment can irreversibly redirect pollen development to embryogenic development. The achievements in this domain have been an improvement of the model system to make it accessible to biochemical and cytological analysis, the isolation of a number of cDNA sequences whose expression is associated with the induction process, the identification of a limited number of unique proteins exclusively synthesized under embryogenic conditions, the identification that rearrangements of the cytoskeleton are associated with the induction process. The onset and timing of DNA, mRNA and protein synthesis were determined in isolated pollen cultured under embryogenic and non-embryogenic conditions. Reentry of the vegetative cell into the cell cycle appeared to be a prerequisite for the induction of embryogenesis. Techniques for the analysis by bidimensional electrophoresis of protein synthesis and phosphorylation during induction of microspore embryogenesis, have been optimized and can now be extented to a variety of studies of cell differentiation. Interesting results concerning the synthesis and histochemical localization of HSP proteins during the induction process have been obtained, raising the working hypothesis that these proteins could be necessary for the initiation of the embryogenic pathway. Some of the proteins and genes associated with the very first stages of plant embryogenesis have been identified and can now be further characterized.

MAJOR SCIENTIFIC BREAKTHROUGHS

Major successes in terms of scientific results have been obtained and new avenues have been opened. In agreement with the terms of the five initial contracts we effectively contributed to improve basic knowledge concerning factors regulating growth and differentiation of plant cells. We also brought totally new results concerning cell division, somatic embryogenesis and microspore embryogenesis. The discovery of the role of secreted proteins and glycoproteins in the the regulation of somatic embryogenesis is a good example of such an important progress. Another exciting example concerns the manipulation of cell numbers per organ resulting from the expression of mutant *cdc2* genes. More generally, the discovery of new genes and gene products involved in several critical and complex aspects of plant differentiation provides unique possibilities to manipulate plant growth and development. This will undoubtedly have important practical implications in the production of transgenic material having improved performance.

Two big T-Projects concerning plants have been launched at the same time: our T-Project and the *Arabidopsis* T-Project. The specific objectives of these research programmes were at the origin apparently quite distinct. In fact, considering the significant title of the *Arabidopsis* T-Project: Molecular Identification of New Plant Genes, the identification of the function of genes involved in the regulation of development has been a common goal for both projects. The strategies used, mutation, gene isolation and mapping on one hand, physiology and biochemistry of protein isolation, gene cloning and transformation on the other hand were largely complementary. As a consequence, *Arabidopsis thaliana* has been rather widely used as a model plant by different groups of our T-Project. We have had a quite significant activity concerning the knowledge of *A. thaliana* development using our biochemical and physiological expertise to isolate and characterize new mutants affected in their response to growth factors, to complement mutants by transformation, to clone new genes involved in signal perception, etc. The list of publications illustrates our contribution to this domain.

The 'production' of a variety of powerful tools is one of the best reflect of the impact of the T-Project and its contribution to the acceleration of research in the plant science community. A variety of nucleic acid probes, mutated genes, gene constructs, transgenic plants, antibodies, oligosaccharides, new cold and labelled ligands, cell lines, etc (see the list in T_{NEWS} No. 13) have been prepared, representing a collection of about 130 new tools. They have been circulated inside the T-Group, raising in 110 cases the opportunity for joint experiments. Importantly, they have also been distributed in 142 cases to laboratories outside the T-Project, illustrating the very positive 'contaminating' aspects of the T-Group activities. The power of these tools is well highlighted by the fact that their use immediately allowed new discoveries in terms of regulation of cell division and differentiation in plants.

As to december 1993, more than 200 papers describing the results obtained in the frame of the T-Project have been published and the list of publications, collected in T_{NEWS} No. 13, will still be enriched during 1994. *The papers published by the group are presented in this volume as five elementary lists in the final report of each of the five contracts.*

Impressive results have been obtained and new and exciting avenues for the future have been opened. Nevertheless, we only scratched the surface of the complex mechanisms underlying the regeneration of plants. Much has still to be done and research in this domain should continue to be supported at an appropriate level taking into account the fact that the amount of money spent for any single animal hormone is larger than the money spent for the whole set of plant growth regulators and that systematic efforts should be pursued on the developmental regulation of the embryogenic process and plant regeneration in general.

MAJOR COOPERATIVE LINKS

This T-Project started on a quite noticeable level of cooperation between the various subgroups and teams involving about 120 scientists. This cooperation has been consolidated all along the three years throughout frequent discussions, considerable exchange of tools and personnel and most importantly by common experiments. Three to six sectorial meetings for each group and three general meetings have been organized where members of different groups gathered and interacted to build fruitful new substates of the initial network. More than 30% of the papers published involved researchers from different laboratories and about 20% corresponded to publications with transnational authorship.

A newsletter, T-News, was created from the very start of the project. It has been, during three years, the internal link between the 25 research teams, but it became rapidly a useful way to diffuse our activities to selected scientists and decision makers, outside the T-Group. Thirteen issues have been circulated. The last two issues, T-News N° 12 and N° 13 give a good idea of the achievements of each team involved in the T-Project (T-News N° 12) and the productivity of the whole group, as illustrated by the list of papers published and tools produced (T-News N° 13).

Both industrial and academic researchers have been involved in this european effort. Four industrial partners have been directly involved in the research activities conducted in the frame of this T-Project. Nickerson Biocem interested in the rapid production of homozygous double haploid plants from microspores has brought an important contribution to the project concerning the microspore embryogenesis by providing a barley microspore culture system in order to allow a comparison with the *Brassica* dicotyledonous system (Contract BIOT CT-900160). Danisco Biotechnology has had a remarkable impact on the research concerning the extracellular proteins secreted by embryogenic carrot cell cultures by providing the technology for large scale culture which allowed the production in mg amounts of the 32 kD chitinase and other extracellular chitinases as well as the isolation of PCR clones for the 32 kD and class IV related chitinases (Contract BIOT CT-900177). This company also contributed to the effort concerning the systematic study of auxin binding proteins by photoaffinity labeling (contract BIOT CT-900178). MOGEN International has developed new microsequencing methods for the analysis of putative receptor proteins and contributed through this effort to the partial sequencing of the ethylene and fusicoccin binding proteins and design of oligonucleotide probes to screen (c)DNA libraries (Contract BIOT CT-900158). LVMH Research contributed to the efforts concerning the use of *rol* genes to modify plant morphogenesis by concentrating on new architectural and flowering phenotypes induced in rose plants by *rolC* expression (Contract BIOT CT-900179).

Structured contacts have also been progressively established with non-contracting industries interested in the activities of our T-Project and of the *Arabidopsis* T-Project. A Plant Industrial Platform (PIP) for both Projects has been created, the idea being that interactions between the PIP and the public laboratories would help to fill the gap between applied and basic research as well as that between model and recalcitrant species. Operational links have been built with mutual exchanges and diffusion of information and attendance to meetings. A PIP Newsletter was created in 1993 and exchanged with our T-News. The Plant Industrial Platform organized an Open Round Table Discussion during the final meeting of the T-Project in Wageningen. Thirteen industrial representatives were present and the discussion, centered on industrial priorities for Plant Science Research in the Framework IV programme, attracted much interest from the members of the T-Project.

Another remarkable level of cooperation has been with the monitoring unit of the T-Project. Such a large and complex project could not have been conducted without the very efficient help of the members of this monitoring unit: Etienne MAGNIEN, who, on behalf of the EC Commission acted as the scientific officer in charge of the implementation of the T-Project at the ECC Biotechnology Division and Pierre PRINTZ, nominated by the BRIDGE Advisory Committee.

European dimension

One feature that distinguishes this EC operation is the impressive large scale network that has been built to promote a multidisciplinary approach through

european cooperation. Most of the subprojects investigated required the cooperation of different disciplines of Plant Sciences, cooperation which could only be found through transnational efforts. Obviously, the multidisciplinary character of this EEC programme has been one of its major strength, providing the adequate range of expertise, instruments and facilities as well as the possibility to use the large variety of tools built up by the group.

Furthermore, special efforts have been devoted by the team of coordinators to facilitate as much as possible the circulation of young scientists attached to the T-Project (attendance to sectorial or general meetings, visits of other european laboratories, short stays in foreign laboratories to undertake common experiments, etc.). A specific round table discussion has been organized during the final meeting in Wageningen by the post doctoral fellows working in the frame of the T-Project. As they represented the major task force of most of the groups, their opinions concerning how the T-Project developed its activities as well as their views concerning future cooperative programmes raised much interest. They expressed their appreciation of the conditions in which they have worked and of the efforts done through TNEWS to give an easy access to the transnational network of research in the domain and to create some sort of 'european' spirit of their activities. According to heterogeneities between laboratories they have been working in, they recommended that all the efforts done to exchange rapidly relevant information, to promote the exchange of young scientists through short stays in different laboratories of the european networks in order to undertake common experiments or to be trained to new techniques should be systematically reinforced in the future cooperative projects. They also expressed a strong concern for the relatively short duration of the contracts they were employed on, calling for more sustained efforts in specific domains of research, associated with longer term contracts and for more information on the possibilities offered by the next european programmes.

Overall, one of the remarkable output of this project has been to build a new scientific community. For european researchers, over the results already obtained, the last three years gave a strong impetus to cooperative work and will have a larger impact for organizing future work in analyzing the complex problem of the regulation of plant development.

PUBLICATIONS: 48 (joint) + 161 (individual)

The publications of the individual subprojects are listed at the end of each corresponding report.

Regulation of the inductive phase of microspore embryogenesis (BIOT CT-900160 — Group 5)

COORDINATOR:
5.1 P. PECHAN, MPI für Biochemie, München, D
PARTICIPANTS:
5.2 A. van LAMMEREN, Agric. Univ. Wageningen, Wageningen, NL
5.3 A. MARTIN, IAPV, Cordoba, E
5.4 A. McQUEEN, Nickersons, Cambridge, GB
5.5 H. DONS, CPRO-DLO, Wageningen, NL

BACKGROUND INFORMATION

The *Brassica napus* microspore culture system proved to be a very good model both as a vehicle for the study of plant cell proliferation and an important tool for the plant breeder. In the normal process of pollen development , the pollen mother cell gives rise to haploid microspores which, after mitosis, mature into pollen grains capable of pollinating flowers. One of the two sperm nuclei fertilizes the egg resulting in a diploid embryo development and a new plant generation. However, by forcing the microspore itself to undergo embryo development, a haploid or diploid plants are formed which are of great interest to plant breeders. A programme of research on the prôduction of double haploids from microspores has been assigned a high priority in many plant breeding companies since the rapid production of homozygous plants is of major commercial value in many crop species.

OBJECTIVES AND PRIMARY APPROACHES

Our aim is to understand the underlying cellular and molecular basis of induction of embryogenesis in plants. To this end we have investigated the cellular and molecular mechanisms associated with the embryogenesis induction process in *Brassica napus* microspores. A heat treatment of 32°C for 8 h is used to induce the process of embryogenesis, and subsequent plant formation, in late uni to early bi-nucleate microspores. From previous studies it is known that such a treatment is sufficient to induce the embryogenic process at high frequencies. We used this knowledge to design experiments whereby the cellular and molecular events associated with the induction process could be clearly identified and investigated. The following main objectives were set:

— isolate and characterize genes and proteins associated with induction of microspore embryogenesis

— characterize the microtubule and microfilament changes associated with the process

— investigate the inheritance of the embryogenic potential within plant populations

— to make available to the Max-Planck Institut (MPI) a highly embryogenic barley (*Hordeum vulgare*) microspore culture system in order to enable a better comparison with the *Brassica napus* microspore culture system.

RESULTS AND DISCUSSION

The results are presented for the participating groups separately.

5.1 Max-Planck-Institut für Biochemie (MPI), Mnchen, Germany

The group investigated the molecular mechanism controlling the induction of microspore embryogenesis in plants using *Brassica napus* microspores as a model system.

Apart from a successful transfer of *Brassica* microspore culture technology to all group 5 members, group 5.1 has concentrated its research efforts on the following areas:

a) Isolation of clones associated with induction of microspore embryogenesis

Selected cDNA clones were partially analyzed for copy number, sequence and expression. 3 categories of induced clones were identified:

- clones containing ribosomal sequences, including mitochondrial rRNAs. They were single to more than 20 copies per genome.
- clones showing no sequence homology to GenEMBL date base. They have 4 and 8 copies per genome respectively.
- clones containing 30-50 bp stretches of simple repeated sequences (SRS) with several copies in the genome.

Northern blots and *in situ* hybridization indicate some of these clones may be specific only to the embryogenic microspores.

Antibodies are currently being raised against these clones.

b) Expression of cell cycle regulated genes in the induction of microspore embryogenesis

A number of probes, derived from genes known to be involved in cell proliferation or clones isolated from microspores induced to undergo the embryogenic process, are currently being tested for their expression and possible importance in the induction of microspore embryogenesis. One such probe, from clone 231, shows specific cellular localization in early bi-cellular pollen.

c) Analysis of gene products.

Total proteins were isolated from *Brassica napus* microspores. The proteins were separated on one and two dimentional gels and analysed: a number of changes in protein profiles were observed during the embryo induction process. For example, Western blot analysis indicates heat shock proteins 68 and 70 are constitutively expressed, whereas heat shock protein 17 is induced only in heat shocked microspores. The localization of these proteins in the cell differs between embryogenic and non-embryogenic microspores. As part of this investigation the *Brassica napus* heat shock 17 gene has been isolated.

The results obtained so far will help us to better understand the regulation of the induction of microspore embryogenesis and, in combination with other techniques, manipulate the inductive process.

5.2. Agricultural University, Wageningen, Holland

The following topics were covered by this group:

a) Microtubular cytoskeleton and the change of position of nuclei in the cells

Microspore and pollen of *Brassica napus* were cultured under conditions leading to embryo formation. Concomitant changes in the microtubular and microfilamental cytoskeletal configurations were analyzed and compared with *in vivo* conditions. The *in vivo* development shows a clear cell cycle dependence. During cultivation two embryogenic pathways started from microspores and were either characterized by turned division planes or by division after the nucleus had migrated towards the cell centre. In both cases microtubules clearly exhibited new arrangements and likely played a major role in newly induced symmetrical division. In pollen, embryogenic development started in the vegetative cell, provided the generative cell was arrested near the pollen wall. The concomitant disappearance of distinct microtubular arrays is likely to be responsible for the positioning of the vegetative cell. The changed position of the nuclei resulted in new division patterns such as symmetrical divisions, and this resulted in embryogenic development. Microfilaments appeared to be of less importance in the change to the embryogenic pathway of development. The microfilamental cytoskeleton

exhibited a loss of polarity in embryogenic cells. However, cytochalasin treatment revealed that microfilaments, do not influence embryogenesis.

b) Immunological analysis

Phosphorylation of proteins was detected by the application of an antibody MPM 2. It was found that significant phosphorylation always occured both in induced and non-induced microspores and pollen. This work was done in collaboration with subgroup 5.5.Immunocytochemical detection of heat shock proteins 68, 70 and 72/73 was also carried out.It was observed that the distribution of HSP 68, 70 and 72/73 at non-embryogenic culture conditions was comparable to that found *in vivo*. HSP 68, restricted to mitochondria, appeared not to change qualitatively by elevated temperatures. HSP 70 and 72/73 clearly showed temperature induced subcellular changes in distribution. These phenomena might be a prerequisite, or might be related to the reentry of repetitive cell cycles which precede embryogenesis. Conclusive experiments are being carried out at present. This work is a collaboration of groups 5.1, 5.2 and 5.5.

c) Changes in the nuclear DNA synthesis

The dynamics of nuclear DNA synthesis were analysed in isolated microspores and pollen of *Brassica napus* that were induced to form embryos. Microspores cultured at non-embryogenic condition (18°C) continued normal gametophytic development. When cultured at embryogenic condition (32°C), they often divided symmetrically and became embryogenic. The labelling pattern revealed that microspores are inducible to form embryos from G1 till G2 phase. Bicellular pollen, cultured at 18°C, exclusively exhibited labelling in generative nuclei, an observation similar to the *in vivo* situation. Early bicellular pollen cultured at 32°C, however, also exhibited DNA replication in vegetative nuclei. The majority of vegetative nuclei reentered cell cycle after 12h of culture, some already after 4h indicating that heat stress evokes an early induction of DNA replication. The DNA replication in the vegetative cells preceded division of the vegetative cell, a prerequisite for pollen derived embryogenesis. At the moment of isolation, microspore nuclei and nuclei of generative cells were at G1, S or G2 phase. Vegetative nuclei of pollen were always in G1 phase at the onset of culture.

5.3. Instituto de Agronomas y Protección Vegetal (IAPV), Cordoba, Spain

The work of the group had three objectives:

- Unravelling the genetic basis of the variability of response for *Brassica napus* microspore culture.
- Mapping the gene/s responsible for those differences
- Producing and characterizing adequate genotypes for every different steps (induction, division, regeneration)of the embryogenic process from microspore to flowering plant.

Three environments, open field , green house and growth cabinet were used to evaluate the induction and process of embryogenesis in microspore culture of *Brassica napus*. Three genotypes, Topaz, Hanna and a rapid cycling cultivar (BN4-22) which differ in culture response and genetic background as well as reproductive cycles were used for the crossing experiments. F1 and F2 generations, backcrosses and parentals of the crosses Topaz x Hanna and BN4-22 x Hanna were evaluated in the three environments. Simultaneously, recombinant inbreed (RI) lines, from each cross, were generated through a single seed descendent program. These RI lines offered advantages over the F2 and backcrosses populations. They may be propagated indefinitely once homozygosity has been attained and increased the probability of recombination of linked genes. Moreover, the heterotic effect of the heterozygosis on F2 plants is bypassed.

When trying to map a quantitative trait loci (QTL), as response to microspore culture in *Brassica napus*, two contradictory requirements had to be fulfilled. First, crosses between parentals differing in many loci were carried out in order to have segregants for many markers. Second, parents differing in the character of study but with similar

genetic background are needed. The cross BN4-22 x Hanna is paradigmatic for the illustration of this point. In order to reduce the life span between generations we used a line with shorter life cycle. This character affected also the inflorescence structure producing less buds per inflorescence, making difficult the isolation of microspores, and increasing the experimental error in regards the analysis of microsporogenesis.

Generations F2, F1, backcrosses and parentals were evaluated for induction, division and regeneration in order to find the genetic system controlling these events. The observed values did not fit a simple model. Individuals with higher and lower expression of the characters were selected on the cross Topaz x Hanna. The evaluation of the progenies of these individuals cover the whole range of variation. It was concluded that the environmental variation (experimental errors included) cover the genetic differences. Nevertheless, in the cross BN4 22 x Hanna the segregation for embryogenic induction/no induction fit a 3:1 ratio when lower than 10% microspore division was considered as non-inductive. Induction was linked to the primer OPA11I at a distance of 20 cM. This linkage is likely an artefact.

Random Amplified Polymorphic DNA (RAPD) markers were used for mapping. From 120 primers tested, 98 (81.66%) give consistent DNA amplification. 77.5% of these products were polimorphic for the two crosses Topaz x Hanna and BN4-22 x Hanna. We have tested many others genotypes, including winter types and alloplasmic lines, and obtained positive response to culture when the plants were grown under favorable conditions. It seems that the observed variability in the response to microspore culture reflects the diferences in traits related to the physiology of flowering or differences in culture media requirements.

Regenerated *Brassica napus* plants from microspores of low embryogenic lines resulted in double haploid (DH) plants with higher frequency than those obtained from high embryogenic lines. The origin of these DH could be spontaneous chromosome doubling or embryogenesis of a non-reduced microspore. The PCR products of the progenies of DH clearly show the homozygotic conditions of DH and the sexual transmission of specific regeneration markers. 85 primers were used and 22 were informative. In microspore regenerated plants the presence of amplification products specific to them (absent on mother plants) have been detected with 12 primers. Crosses between donor plants and their DH have been performed and F1, F2 and backcrosses analyzed to uncover the genetic basis of these amplified products. Up to now, only 15 F2 and 20 backcrosses plants have been analyzed, but the genetic system controlling these markers is not simple, in some cases more than one locus is involved. In two DH analysed similar genetic sequences have been amplified suggesting specificity of the DNA amplification as a result of microspore culture. Amplified chromosome segments or chromosome rearrangements are well documented in tissue culture due to somaclonal variation.

In order to diminish the variability included by the isolation procedure, in which many flower were used in each isolation, we cultured individual anthers, where moreover the stage of development was assesed. With this procedure we did not found non-inducible genotypes. Moreover, a Hanna genotype which previously we defined as inducible but non divisible (non regenerative) with this procedure turned out to be high embryogenic.

5.4 Nickerson BIOCEM, Cambridge, England

Experiments were carried out to characterise the highly embryogenic monocot barley (*Hordeum vulgare*) microspore culture system in order to enable a better comparison with the already well characterised dicot *Brassica napus* microspore culture system.

a) Culture conditions

Embryogenic divisions were observed after 48 — 72 hours culture at 25°C (permissive temperature). Lower culture temperatures were tested to define a non-permissive temperature to parallel the *Brassica napus* system. However, high frequencies of embryogenesis were observed at 18°C and 12°C.

b) Developmental stage

In common with *Brassica napus*, uninucleate microspores and binucleate microspores (pollen) may be induced to undergo embryogenesis. In our experiments spikes containing microspores at the mid-uninucleate stage were subjected to a cold treatment of 4°C for 28 days prior to mechanical isolation of the microspores. DAPI staining revealed that, following cold pretreatment, the majority of isolated microspores were binucleate. An alternative pretreatment involved culturing excised anthers on 0.3 M mannitol for 4 days. Microspores isolated from these anthers were still uninucleate which may have significance for transformation experiments. However, isolations following mannitol treatment were found to be less embryogenic than those following cold treatment.

c) Cell types

Three different types of cell could be distinguished in the initial stages of microspore culture. One type consisted of small, plasmolysed microspores which were apparently dead. A second type consisted of swollen microspores, rich in granular cytoplasm. A third type also consisted of swollen microspores, but possessing a large vacuole traversed by cytoplasmic strands. Intermediate types were also observed. It is proposed that the swollen, granular type divide and grow while the swollen, vacuolar type die.

Microspore cultures which had been subjected to various pretreatments and culture conditions were supplied to the MPI for RNA extraction and probing.

Nickerson BIOCEM will continue to co-operate with the MPI in assessing the results of the project and exploring their applications for tissue culture and for plant breeding.

5.5 Centre for Plant Breeding and Reproduction Research (CPRO-DLO), Wageningen, Holland

The work of the group was concerned with the identification of proteins that might be involved in the induction of embryogenic development of *Brassica napus* microspores and is carried out in collaboration between groups 5.1, 5.2 and 5.5.

a) Two dimentional (2-D) gel electrophoresis was used to analyse the [^{35}S]-methionine incorporation in proteins during the first 8 hours under embryogenic conditions at 32°C. Qualitative and quantitative computer analyses of 2-D [^{35}S]-methionine protein patterns revealed six polypeptides specifically labeled under embryogenic culture conditions. Eighteen polypeptides incorporated [^{35}S]-methionine at a significantly higher rate under embryogenic culture conditions (32°C) than under non-embryogenic culture conditions (18°C), whereas one protein was preferentially labeled at (18°C). None of these proteins could be labeled with [^{32}P]-orthophosphate in situ during the same 8h labeling period. These results indicate that only a limited number of proteins detectable in the 2-D gels of microspore extracts are associated with the early induction of embryogenesis.

It was shown that the induction of microspore embryogenesis is associated with rearrangements of the microtubular cytoskeleton (subgroup 5.2). Immunoblotting was used to see whether these cytological changes were accompanied by differential synthesis of the cytoskeletal proteins. Multiple isoforms of β-tubulin, ß-tubulin, and actin were detected in both embryogenic and non-embryogenic microspore cultures, but no differential synthesis of any of these isoforms was observed during the first 8h of embryo induction.
The reproducible identification of the differentially radiolabeled proteins in the 2-D gels will allow the sequencing of representative peptides and the isolation of the corresponding cDNAs. This may lead to the identification and characterization of proteins associated with the very first stages of plant embryogenesis.

b) The database of [^{35}S]-methionine labeled proteins was also used to study the synthesis and localization of a number of heat shock proteins from the 70 kDa class during the induction of microspore embryogenesis. Heat shock proteins were identified by Western-blotting of two-dimensional (2-D) gels from microspore protein extracts

derived from embryogenic cultures. Three antibodies were used: a polyclonal antibody raised against tomato HSP70 (anti-HSP70), a monoclonal antibody raised against human HSP72/73 (anti-HSP72/73), and a polyclonal antibody raised against tomato HSP68 (anti-HSP68). The immuno-reactive protein spots were matched to the [^{35}S]-methionine labeled proteins. It was found that several of the proteins that were preferentially synthesized in the embryogenic cultures belonged to the 70 kDa HSP class.

MAJOR SCIENTIFIC BREAKTHROUGHS

Important findings:

- improved definition of the microspore inductive phase: DNA synthesis begins within 4 hours of microspore culture initiation and first embryogenic division is associated with changes in the microtubule orientation.
- Specific localization of heat shock proteins (Hsp 17 and Hsp 70) in embryogenic microspores is associated with the inductive process.
- demonstrated *in vivo* phosphorylation of a HSP70-like protein, associated with induction of microspore embryogenesis.
- isolated a number of cDNA clones which are associated with induction of microspore embryogenesis

Techniques:

- a promising start was made with development of a microspore embryogenesis system for *Arabidopsis*.
- a reproducible and sensitive system has been developed for comparative analysis of protein synthesis during the first 8 hours of microspore culture.

The rapid production of homozygous, double haploid plants from microspores is of major commercial value in many breeding programmes since double haploids permit the efficient selection of desirable gene combinations and enable new traits to be incorporated into product lines or hybrids more quickly. An improved understanding of the induction of microspore embryogenesis would open the possibility to manipulate and transfer the genes regulating this process.

MAJOR COOPERATIVE LINKS

Throughout the three years a number of meetings were held between the participating laboratories.

There have been frequent meetings between MPI, Agric. Univ. and CPRO-DLO (groups 5.1, 5.2 and 5.5). Apart from a number of brief visits, two joint meetings were held in 1992 and 1993. Both dealt with interpretation of results and exchange of probes for further research activities. The cooperation with MPI focused on the application and analysis of embryogenesis specific mRNA probes, *in situ* and heat shock protein analysis. Cooperation with CPRO-DLO concerns research on phosphorylation events during the induction of embryogenesis. The biochemical aspect is treated by CPRO-DLO, the structural component by Agric. University.

The group of Agric. Univ. had cooperation with other groups of the T-Project: V. Onckelen (3.4), Palme (1.1), Venis (1.2) and Klambt (2.2). These cooperations concerned the in situ detection of plant growth regulators and their binding proteins or receptors.

Bilateral visits also took place also between MPI,CSIC and Nickersons (groups 5.1, 5.3 and 5.4). The meeting in Spain dealt with solving practical problems associated with high temperatures during the *Brassica* growing season in Spain. The meeting with Nickersons was to ascertain the progress of *Brassica* and barley microspore research.

All laboratories in the T-project participated at three joint meetings. The first was held in Köln in 1991 and the subsequent two in Wageningen in 1992 and 1993. Apart from

the presentation of current results, discussions took place on future interactions between subgroups, such as joint publications, common use of probes for in situ hybridization, exchange of *Brassica* lines and cooperation on 2D gel analysis. Indeed all these interactions were strengthened and expanded throughout the year.

Additionally, cooperation with Czech colleagues was carried out at all participating groups through the TEMPUS project, for example on the quantification of DNA contents of nuclei and mRNA analysis.

Because of the subject nature of the research, it is by necessity a transnational collaboration. The group works well together and thus allows quicker transfer of knowledge and technology within the EC and beyond. The personal contacts have meant improved confidence and willingness for future cooperation. This is already reflected in the fact that 4 of the group members participate in two other EC projects, including student exchanges.

PUBLICATIONS

Joint publications

Hause, B., Hause, G., Pechan, P., van Lammeren, A.A.M. (1993) Cytoskeletal changes and induction of embryogenesis in microspore and pollen cultures of *Brassica napus* L. Cell Biology Intern. **17(2)**: 153-168

Cordewener,J., Hause,G., Gorgen, E., Businik, R., Hause, B., Dons, H.,van Lammeren A., van lookeren Campagne, M., and Pechan P. Changes in the synthesis and location of heat shock proteins in embryogenic microspores of B. napus (submitted).

Gorgen. E, Neumann, D., Cordewener J.G.W., Honys,D. and Pechan, P. Localization of HSP 17 in microspores is associted with induction of the embryogenic process (1994) (submitted)

Angelis, K., Havlicky, J., Hause, G., van Lammeren, A. and Pechan, P. Isolation and characterization of a microspore specific gene (submitted)

Individual publications

Binarova, P., Straatman, K., Hause, B., Hause, G., van Lammeren, A.A.M. (1993) Nuclear DNA synthesis during the induction of embryogenesis in cultured microspores and pollen of *Brassica napus* L. TAG **87**: 9-16.

Custers., J.B.M. J.H.G. Cordewener, Y. Nollen & J.J.M. Dons and Van Lookeren Campagne (1993) Temperature controls both the sporophytic and gametophytic development in isolated microspore cultures of *Brassica napus*. Plant Cell Reports (in press).

Cordewener J.G.H. Businik, R., Traas, J.A. Custers J.B.W., DonsJ.J.M. and Van Lookeren Campagne (1993) Induction of microspore embryogenesis in Brassica napus is accompanied by specific changes in protein synthesis. Planta (in press).

Pechan, P.M. (1991) Heat shock proteins and cell proliferation FEBS Lett. **280**, 1-4

Pechan, P., Bartels, D., Brown, D.C.W., and Schell, J. (1991) Messenger RNA and protein changes associated with induction of microspore embryogenesis. Planta **184**: 161-165

Duijs, J.G., R.E. Voorrips, D.L Visser & J.B.M. Custers, (1992). Microspore culture is successful in most types of *Brassica oleracea L.* Euphytica **60**: 45-55.

Traas J.A. A.F. Beven, J.H. Doonan, J.H.G. Cordewener & P.J. Shaw, (1992). Cell cycle dependent changes in labelling of specific phosphoproteins by the monoclonal antibody MPM-2 in plant cells. The Plant Journal **2**: 723-732.

de Jong, A, J. Cordewener, F. Lo Schiavo, M. Terzi, J. Vandekerckhove, A. van Kammen & S. de Vries, (1992). A carrot somatic embryo mutant is rescued by chitinase. The Plant Cell **4**: 425-433.

Hause, G., Hause, B., van Lammeren, A.A.M. (1992) Microtubular and actin filament configurations during microspore and pollen development in *Brassica napus* cv. Topas. Can J Bot. **70**: 1369-1376

Hause, G., B. Hause and A.A.M. Van Lammeren (1992). Changes in the microtubular cytoskeleton during microspore and pollen development in *Brassica napus* L. Acta Bot. Neerl. **41**, 221.

Angelis, K.J. and P. Pechan (1993). Analysis of stress induced embryogenesis in *Brassica napus*. J. Cell Biochemistry, Supplement **17B**, 26

Brown, D.C.W., E.M.Watson, and P. Pechan (1993) Induction of desiccation tolerance in microspore-derived embryos of *Brassica napus*. In Vitro, **29**: 113-118.

Plant growth regulators: perception, interaction and response (BIOT CT-900158)

COORDINATOR:
M.A. HALL, Institute Biol. Sciences, University of Wales, Aberystwyth, GB
PARTICIPANTS:
K.R. LIBBENGA, University of Leiden, Leiden, NL
D. KLÄMBT, University of Bonn, Bonn, D
P. ADUCCI, University of Rome 'Tor Vergata', Rome, I
Mogen Int. NV, Leiden, NL

BACKGROUND INFORMATION

The regeneration of whole plants from tissue cultures, organ cultures and so on, is vital to efforts aimed at crop improvement. However, although such regeneration is relatively straightforward for many species, this is not true for many others, including a large number with agricultural and horticultural importance. It has been known for many years that an understanding of the mechanism(s) of action of plant growth regulators is central to a solution of the problem of recalcitrant species and the project outlined here sought to deepen and extend our knowledge of the means whereby plants perceive growth regulators (through receptors) and how receptor/ligand interactions are transduced to modulate developmental responses. The studies were largely at the biochemical and molecular level.

OBJECTIVES AND PRIMARY APPROACHES

The objectives of the study were several-fold and included the isolation and characterisation of the receptors for ethylene and fusicoccin and the cloning of the appropriate genes, together with work on auxin receptors other than that already isolated from maize. Parallel studies sought to determine receptor functionality using a variety of strategies. A major part of the work was both to characterise transduction pathways and their interactions including work with protoplasts and anti-receptor antibodies or impermeant hormone analogues, *in vitro* reconstitution of components of perception and transduction chains and the use of *Arabidopsis* sensitivity mutants.

A further dimension of the work was to elucidate further, particular transduction mechanisms for auxin, ethylene and fusicoccin, using paradigms from animal systems.

RESULTS AND DISCUSSION

By the use of a diverse array of protein purification techniques, fusicoccin receptors have been successfully purified from the plasmalemma of maize shoots. Three bands have been observed on SDS-PAGE at 31, 33 and 90 kDa The use of photoaffinity labelling indicates that the functional receptor is the 90 kDa band. Sequencing of the 30 kDa area reveals complete homology with the 14-3-3 class of proteins which are widely distributed in eukaryotes and which, *inter alia*, are known to regulate protein kinase C-like kinases, indicating a role in the early events of fusicoccin signal transduction. Polyclonal antibodies raised to a sequenced 20 amino acid peptide from the 30 kDa area recognise a similar protein in partially purified fusicoccin receptor preparations but do not cross-react with the 90 kDa band.

The ethylene receptor is a highly hydrophobic protein, which complicated efforts to purify it. However, a combination of FPLC separations and the development of a technique of 'semi-denaturing' electrophoresis has resulted in the characterisation of two bands on gels at 28 and 26 kDa which bind ethylene and which can further be uniquely separated by two dimensional electrophoresis. N-terminal and internal sequences have been obtained for the 28 kDa band; there is no significant homology with any protein in the data banks.

Specific antibodies have been raised to both components. There is cross-reactivity with both 26 and 28 kDa bands but the signal is much stronger with the homologous components. The antibodies were used to screen cDNA expression libraries from *Arabidopsis* leaf tissue and from *Phaseolus* abscission zones. This approach was unsuccessful for such low abundance proteins, hence oligonuceoltide probes derived from the amino acid sequences were used for library screening and PCR studies. This work is ongoing and it is hoped that these approaches will provide information concerning the gene(s) for protein(s) involved in ethylene perception.

The antibodies referred to above have also been used to probe for homologous proteins in other species. These have been detected in peas, *Arabidopsis* and rice. Equally, immunological studies indicate that ethylene-sensitive tissues (e.g. abscission zones in *Phaseolus*) possess higher concentrations of antigen than less sensitive tissue; observations confirmed by direct measurement of ethylene binding.

New techniques have been developed for identifying and characterising ethylene binding sites *in vivo*. These techniques have revealed the presence of ethylene binding sites in all tissues examined (*Arabidopsis*, tomato, rice and peas). Moreover, this work has shown that there exist at least two classes of site. These are identical in all respects (e.g. affinity, specificity) save that one class shows fast association/dissociation kinetics and the other slow association/dissociation kinetics. Further work indicated that while the slow-associating components are membrane-associated and analogous to the sites in *Phaseolus* the fast-associating sites are either membrane-associated of cystolic.

Studies have been carried out an a range of ethylene sensitive mutants of *Arabidopsis* (the *eti* series). The mutations are pleiotropic suggesting that they affect ethylene perception or transduction events unique to the growth regulator. In common with other species, *Arabidopsis* possesses both fast and slow-associating ethylene binding sites and the work indicates that differences in sensitivity may be associated with changes in the abundance of either class. Thus, the mutant showing complete insensitivity to ethylene (*eti* 5) shows a significant reduction in abundance of fast-associating sites compared to wild type; on the other hand *eti* 13 — which is the second most insensitive mutant — shows normal concentrations of fast-associating sites, but appears to lack the slow-associating variety.

The use of antibodies raised to maize auxin binding protein has enabled the cloning of three homologous genes from rice. A range of antibodies have been raised to a number of epitopes on maize auxin binding which have either auxin agonist or antagonist activity.

It was shown in tobacco cell suspension cultures that certain genes belonging to related families (the Nt series) are induced by auxin treatment of hormone-starved cells. Proteins encoded by the genes were found to share a limited but significant homology with glutathione-S-transferase (GST) and proteins produced by such genes in transformed *E. coli* have GST activity. The promoters of such genes contain information to direct root-tip specific and auxin-inducible expression *in planta*.

Similar experiments with Nt107 and Nt114 genes also showed that an auxin-responsive element (AURE) resides in the promoter region. Further analysis of two promoters of each gene family led to the identification of an As1-like element (TGACG motif) located around position — 300 prior to the ATG codon, probably likely represents the AURE. Gel retardation assays of promoter fragments showed that a nuclear protein related to ASF-1 (CREB-like, identified as binding to TGACG motifs in the CaMV 35S promoter) binds specifically to the As1-like elements of the Nt promoters although the sequences of the binding site are rather divergent from the site present in the 35S promoter. Synthetic 20 bp DNA sequences representing the As1-like elements were shown to be able to confer inducibility to a heterologous minimal promoter by using a transient expression system employing particle bombardment.

Expression of sense and antisense Nt103 mRNA demonstrated that the presence or absence of this mRNA alone is not crucial for cell division. During these studies it was found that non-auxins, like salicyclic acid (SA) and yeast extract (YE) were also potent inducers of Nt103 genes. A protoplast system was developed using cells which contained a Nt103 promoter fused to GUS. The effect of D16 antibodies (obtained from Prof. Venis, subgroup 1.2) directed against the presumed auxin-binding site of the major maize auxin-binding protein (ABP), which were shown to have auxin-agonist activity, as they induce hyperpolarisation in tobacco mesophyll protoplasts, was investigated in this system. The D16-induced hyperpolarisation was present. It appeared however, that auxin-induced expression of the reporter gene was blocked, while the induction by SA and YE was unhampered. Apparently auxin has (at least partially) a different signal transduction pathway than the other inducers. Screening a cDNA expression library from tobacco cells with D16 antibodies yielded chaperonin-like proteins and a less abundant protein of about 120kD. An ABP homologue was not detected.

Protoplasts from tobacco cell cultures were investigated using the patch-clamp technique and intracellular microelectrode measurements. By means of the patch-clamp technique (whole cell configuration) both an outward and an inward rectifying K^+ conductance could be found in protoplasts from stationary cells, although the inward rectifying conductance was only present in 30% of the protoplasts tested. These conductances were characterised in detail. The studies were extended to *Arabidopsis thaliana* mesophyll protoplasts. In some protoplasts at least two other currents could be found beside the tobacco-like K^+ conductances. In order to make plant cells accessible to patch clamp measurements without the use of cell wall degrading enzymes, a method to employ laser microsurgery was developed in co-operation with Dr. A. H. de Boer (Free University, Amsterdam).

The transduction chain for the fusicoccin signal has been investigated by a dual reconstituted system where the pure plasmalemma H^+-ATPase has been inserted into liposomes with crude solubilised fusicoccin receptors. Optimal conditions for reproducible reconstitution and the biochemical characterisation of the fusicoccin effect on this system have been studied. Recently, the approach of dual reconstitution was used to insert into liposomes fusicoccin receptors from maize shoots and the H^+-ATPase from *Neurospora crassa*, an organism usually insensitive to fusicoccin. The H^+-pumping activity of this enzyme was activated by fusicoccin only in the presence of fusicoccin receptors, suggesting a more general regulatory role for these proteins on the activity of plasmalemma H^+-ATPases. They probably act by affecting the regulatory C-terminus domain of H^+-ATPases. The results obtained with this heterologous system indicate that fusicoccin can be considered as a tool

not only in the study of hormone mode action but also in the elucidation of plasmalemma H^+-ATPase regulation.

The possible role of the phosphoinositol cycle and of G-proteins has been investigated in fusicoccin signalling. As far as can be determined, no role seems to be played by G-proteins in fusicoccin signal transduction. On the other hand, data obtained in *in vivo* experiments indicate that fusicoccin can promote the enhancement of IP3 levels in maize coleoptiles.

It was demonstrated that, in peas, ethylene enhances phosphorylation of specific proteins. The effect is concentration-dependent and is specifically inhibited by the ethylene antagonist norbornadiene. The effect is detectable both *in vivo* and *in vitro*. Work with cell-free preparations indicates that under conditions where phosphorylation is promoted, ethylene binding is inhibited, whereas the reverse is true when dephosphorylation is enhanced. It appears that the proteins, phosphorylation of which is enhanced by ethylene, include those homolgous with the ethylene binding proteins from *Phaseolus*. This observation has been confirmed by immunoprecipitationn using the antibodies raised to the *Phaseolus* proteins. The situation is complicated by the presence of homologous cytosolic proteins which are also phosphorylated. These observations suggest a possible mechanism for the control of receptor activity via phosphorylation/dephosphorylation analogous to animal paradigms such as epidermal growth factor.

MAJOR SCIENTIFIC BREAKTHROUGHS

- Purification, characterisation and sequencing of the fusicoccin receptor and of a 30 kDa protein homologous with 14-3-3 proteins.
- Purification, characterisation and sequencing of the ethylene receptor.
- Identification of auxin-induced genes as glutathione-S-transferase homologues.
- Mapping of the auxin-responsive elements in the auxin-induced GST genes of tobacco.
- Characterisation of the major ion currents in tobacco and *Arabidopsis* protoplasts.
- Demonstration of the involvement of protein phosphorylation in the mediation of the ethylene response and of the inositol phosphate pathway in the transduction of the fusicoccin response.

MAJOR COOPERATIVE LINKS

There have been extensive exchanges of personnel both short and long term, between all groups within this group, with other groups within the T-Project (e.g. Horticulture Research International and Leiden, Leiden and Gif-sur-Yvette) and with groups elsewhere in the EU (e.g. Aberystwyth and Antwerp) and indeed with Eastern Europe (e.g. Bulgaria), and the Former Soviet Union (Moscow).

The emphasis has been always to encourage the exchange of young research workers.

PUBLICATIONS

Joint publication

Mennes A. M., Maan A., Hall M. A. Plant Hormone Receptors. In: Cell to Cell Signals in Plants and animals: Progress Report (eds V. Neuhoff, J. Friend), NATO/ASI Series H: Cell Biology, vol. 51, Springer Verlag, pp 301-314. (1991).

Individual publications

Aducci P., Ballio A., Ferlito B., Fullone M. R., Marra M. Biochemical characterisation of a dual reconstituted system responsive to fusicoccin. 14th International Conference on Plant Growth Substances, Amsterdam, (1991). p.103.

Aducci P. Fusicoccin receptors: perception and transduction of fusicoccin signal. International Symposium 'Biochemical Mechanisms involved in growth regulation'. Milano, (1991).

Aducci P., Ballio A., Fullone M. R., Marra M. Study of fusicoccin sigalling by a reconstituted system. International Symposium on 'Adenine Nucleotides in Cellular Energy Transfer and Signal Transduction' (1991).

Aducci P., Ballio A., Fullone M. R., Marra M. Membrane receptors for fusicoccin: a molecular study. 12th Long Ashton International Symposium 'Transport and receptor proteins of plant membranes: molecular structure and function. (1991).

Aducci P., Benoit C., Fullone M. R., Marra M. Characterisation of GTP-binding proteins from maize. Giorn. Bot. Ital. 125, n.4. p.499 (1991).

Aducci P., Ballio A., Fogliano V., Fullone M. R., Marra M. Functional reconstitution of purified fusicoccin receptors. VII Convegno Nazionale Protein 92, Pavia, p. 26. (1992)

Aducci P. Perception and transduction of hormone signals in higher plants 8th Lecture Course on Biophysics and Molecular Biology 'Molecular Biology of Plants' (Italian Society of Biophysics and Molecular Biology) Cividale del Friuli (Udine), 5- (1992).

Aducci P., Ballio A., Fogliano V., Fullone M. R., Marra M. Molecular studies on fusicoccin receptors and their relationships with H^+-ATPase. 9th International Workshop on Plant Membrane Biology, Monterey, California, USA, (1992).

Aducci P., Ballio A., Fogliano V., Fullone M. R. Marra M., Proietti N. Photoaffinity labelling of fusicoccin receptors in maize. Italian Biochemical Society Transactions, 3 p. 442, **190**, (1992)

Aducci P., Ballio., Battirossi P., Donini V., Fogliano V., Fullone M. R., Marra M., Proietti N. Aspetti molecolari del meccanismo d'azione della fusicoccina Petria 2, **17**. (1992)

Aducci P., Marra M., Mattei M., Nasta D., Proietti N. Messa a punto di un protocollo per la produzione di anticorpi anti-idiotipo verso il recettore per la fusicoccina XXXII Congresso Societa Italiana di fisiologia Vegetale Udine No. 28. (1992).

Aducci P., Ballio A., Fogliano V., Fullone M. R. Marra M. Proietti N. Marcatura del recettore della fusicoccina con un ligando fotoaffine XXXII Congresso Societa Italiana de Fisiologia Vegetale Udine No. 114. (1992)

Aducci P., Ballio A., Fullone M. R., Marra M. Biochemical characterisation of a dual reconstituted system responsive to fusicoccin. In: 'Progress in Plant Growth Regulation', Karssen C. M., Van Loon L. C., Vreugdenhil D., Eds, Kluwer Acad. Publishers, 686-690, Dordrecht, The Netherlands. (1992)

Aducci P., Ballio A., Fullone M. R., Marra M. Fusicoccin receptors: perception and transduction of fusicoccin signal. In: 'Biochemical mechanisms involved in growth regulation', Smith C. J. and Gallon J. Eds, Oxford Univer. Press, (in press). Oxford, Regno Unito. (1992)

Aducci P., Ballio A., Fogliano V., Fullone M. R., Marra M., Verzili D. Membrane receptors for fusicoccin: a molecular study In: 'Transport and receptor proteins of plant membranes', Clarkson D. T., Cooke D. T., Eds, Plenum Publishing Corporation, pp. 179-184. New York USA. (1992).

Aducci P., Cozzella M. L., Di Giorgio D., Fogliano V. and Marra M. Cercospora beticola toxin: a reassessment of some *in vitro* effects. Plant Science 84, 53-57. (1992).

Aducci P., Ballio A., Fogliano V., Fullone M. R., Marra M., Proietti N. Purification and photoaffinity labelling of fusicoccin receptors from maize. European Journal of Biochemistry, **214**, 339-445. (1992).

Aducci P. Perception and transduction of hormone signals in higher plants. Giornale Botanico Italiano, **126**, 921-931. (1992).

Aducci P., Ballio A., Fogliano V., Fullone M. R., Marra M., Verzili D. Membrane receptors for fusicoccin: a molecular study In: 'Transport and receptor proteins of plant membranes', Clarkson D. T., Cooke D. T., Eds, Plenum Publishing Corp. 179-184. (1992).

Aducci P., Ballio A., Fullone M. R., Marra M. Biochemical characterisation of a dual reconstituted system responsive to fusicoccin. In: 'Progress in Plant Growth Regulation', Karssen C. M., van Loon L. C., Vreugdenhill D. Eds Kluwer Publishers, pp. 686-690. (1992).

Aducci P., Cozzella M. L., Di Giorgio, Fogliano V., Marra M. Cercospora beticola toxin: a reassessment of some *in vitro* effects. Plant Science, **84**, 53-57. (1992).

Aducci P. Recettori per la fusicoccina: un approcio molecolare allo studio della percezione transmissione del segnale nelle piante. Scuola Nazionale 'Metodologie avanzate in fisiologia vegetale'. Societa Italiana di Fisiologia Vegetale. Bressanone. (1993).
Aducci P. and Ballio A. Receptors for fusicoccin higher plants. XV International Botanical Congress, Abstracts p.148 Yokohama, Japan. (1993).
Aducci P., Ballio A., Donini V., Fogliano V., Fullone M. R., Marra M. Phospholipase A_2 affects the activity of fusicoccin receptors. FEBS Letters, **320**, 173-176. (1993).
Aducci P., Ballio A., Fogliano V., Fullone M. R., Marra M., Proietti N. Purification and photoaffinity labelling of fusicoccin receptors from maize. Eur. J. Biochem., **214**, 339-345. (1993).
Aducci P., Ballio A., Fullone M. R., Marra M. Fusicoccin receptors: Perception and transduction of fusicoccin signals In: 'Biochemical Mechanisms Involved in Growth Regulation', Smith C. J., Eds, Oxford University Press, 273-279. (1993).
Aducci P., Ballio A., Donini V., Fogliano V., Fullone M. R. and Marra M. Phospholipase A_2 affects the activity of fusicoccin receptors. Federation of European Biochemical Societies, **320** 173-176. (1993).
Ballio A. and Aducci P. Fusicoccin receptors: mode of action of a microbial phytotoxin. In 'Macromolecules in the functioning cell', Proceedings of the 6th Soviet-Italian Symposium, Leningrad, Bayev A. A., Mirzabekov A. D., Timofeeva M. Ya., Eds Puschino. pp. 338-348 (Puschino, Russia). (1991)
Ballio A. and Aducci P. The fusicoccin receptor, a ubiquitous protein of higher plants. In: 'Macromolecules in the functioning cell', Proceedings of the 7th Soviet-Italian Symposium Taormina, 24-27/10/1990, Volpe P., Gambria A., Macaione S., Eds, Pacini, 131-133, Pisa Italy. (1991)
Ballio A. and Aducci P. Fusicoccin signalling in plants. XV International Botanical Congress, Abstracts p. 113. Yokohama, Japan. (1993).
Brailsford R., Voesenek L. A. C. J., Blom C. W. P. M., Smith A. R., Hall M. A., Jackson M. B. Enhanced ethylene production by primary roots of *Zea mays L.* in response to sub-ambient partial pressures of oxygen. In: 'Cellular and Molecular Aspects of the Plant Hormone Ethylene', Kluwer Academic Publishers, pp. 369-370. (1993).
Boot K. J. M., Van der Zaal B. J., Velterop J., Quint A., Mennes A. M., Hooykaas P. J. J., Libbenga K. R. Further characterisation of expression of auxin-induced genes in tobacco (*Nicotiana tabacum*) cell-suspension cultures. Plant Physiol, **102**, 513-520. (1993).
Cozzella M. L., Aducci P., Ballio A., Gasparrini F., Misiti D. Inhibition of fusicoccin-stimulated proton extrusion by maize root extracts. International Symposium 'Biochemical Mechanisms involved in growth regulation'. (1991).
Droog B., Hooykaas P. J. J., Libbenga K. R., Van Der Zaal E. J. Proteins encoded by an auxin-regulated gene family of tobacco share limited but significant homology with glutathione s-transferase and one member indeed shows *in vitro* gst. activity. Plant Mol. Biol. (1993), (in press).
Fullone M. R., Aducci P., Ballio A., Fogliano V., Marra M. Purification of fusicoccin receptors from maize shoots. Giorn Bot. Ital. 125, No. 4. p. 506. (1991)
Hall M. A., Berry A. W., Cowan D. S., Evans J. S., Harpham N. V. J., Moshkov, I. Ye., Novikova G., Raskin I., Smith A. R., Turner R. J., Xiuqing Z. Ethylene recepetors. In: 'Biochemical Mechanisms Involved in Growth Regulation', (C. J. Smith, J. Gallon, eds) Oxford University Press, (1992) (in press).
Hall M. A., Berry A. W., Harpham N. V. J., Roveda-Hoyos G., Smith A. R., El-Abd S. O. Stress physiology in the context of protected cultivation. Acta Horticulurae, **323**, 379-400. (1993).
Hall M. A., Aho H. M., Berry A. W., Cowan D. S., Harpham N. V. J., Holland M. G., Moshkov, I. Ye., Novikova G., Smith A. R. Ethylene recepetors. In: 'Cellular and Molecular Aspects of the Plant Hormone Ethylene'. (J. Pech, A. Latche, C. Balague, eds). Kluwer Academic Publishers, pp. 168-173. (1993).
Harpham N. V. J., Berry A. W., Knee E. M., Roveda-Hoyos G., Raskin I., Sanders I. O., Smith A. R., Wood C. K., Hall M. A. The effect of ethylene on the growth of development of wild type and mutant *Arabidopsis thaliana* (L). Heynh. Ann. Bot., **68**, 55-61. (1991)
Klambt D. NPA inhibits secretion of amylase by barley aleurone cells and auxins can overcome this inhibition. Bot. Acta. **106**, 52-57. (1993)

Klambt D., Knauth B., Dittmann J. Auxin dependent growth of rhizoids of *Chara globularis*. Physiol. Plant. **85**, 537-540. (1992)
Marra M., Aducci P. Ballio A., Fullone M. R. Properties of liposome-reconstituted H^+-ATPase and fusicoccin receptors from maize. Giorn. Bot Ital. 125, No. 4. p. 512 (1991).
Marra M., Ballio A., Fullone M.R. and Aducci P. Some properties of a functional reconstituted plasmalemma H^+-ATPase activated by fusicoccin. Plant Physiology 98, 1029-1034, Baltimore, Maryland, USA. (1992).
Marra M., Ballio A., Fullone M. R., Aducci P. Some properties of a functional reconstituted plasmalemma H^+-ATPase activated by fusicoccin. Plant Physiol., **98**, 1029-1034. (1992)
Marra M., Ballio A., Battirossi P., Fogliano V., Fullone M. R., Slayman C. L. and Aducci P. The fungal H^+-ATPase from *Neurospora crassa* reconstituted with fusicoccin receptors senses fusicoccin signal. (submitted).
Marra M., Fullone M. R. Fogliano V., Pen J., Mattei M., Masi S and Aducci P. The 30 kDa protein present in purified fusicoccin receptor preparations is a 14-3-3 like protein. (submitted).
Mennes A. M., Quint A., Gribnau J. H., Boot C. J. M., Van Der Zaal E. J., Maan A. C., Libbenga K. R. Specific transcription and reinitiation of 2, 4-D-induced genes in Tobacco nuclei. Plant Mol. Biol., **18**, 109-117. (1992).
Moshkov I. Ye., Novikova G. V., Smith A. R., Hall M. A. *In vitro* study of ethylene binding sites in pea seedlings. In: 'Cellular and Molecular Aspects of the Plant Hormone Ethylene', Kluwer Academic Publishers, pp. 195-196. (1993).
Novikova G. V., Moshkov I. Ye., Smith A. R., Hall M. A. Ethylene and phosphorylation of pea epicotyl proteins. In: 'Cellular and Molecular Aspects of the Plant Hormone Ethylene', Kluwer Academic Publishers, pp. 371-372. (1993).
Sanders I. O., Ishizawa J., Smith A. R. & Hall M. A. Ethylene binding and action in rice seedlings. Plant Cell Physiol. **30**, 1091-1099. (1991).
Sanders I. O., Harpham N. V. J., Raskin I. Smith A. R., Hall M. A. Ethylene binding in wild type and mutant *Arabidopsis thaliana* (L.) Heynh. Annals. Bot., **68**, 97-104. (1991).
Sanders I. O., Smith A. R., Hall M. A. Ethylene binding in epicotyls of *Pisum sativum L.* cv Alaska. Plants, **183**, 209-217. (1991).
Van Der Zaal E. J., Droog F. N. J., Boot C. J. M., Hensgens L. A. M., Hoge J. H. C., Schilperoot R. A., Libbenga K. R. Promoters of auxin-induced genes from tobacco can lead to auxin-inducible and root tip-specific expression. Plant Mol. Biol., **16**, 983-998. (1991).
Van Duijn B., Ypey, D. L., Van der Molen L. G., Libbenga K. R. Whole-cell patch-clamp experiments on tobacco protoplasts. In: 'Progress in Plant Growth Regulation', (C. M. Karssen, L. C. van Loon, D. Vreugdenhill, eds.) Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 668-674. (1992).
Van Duijn B. Hodgkin-Huxley analysis of whole-cell outward rectifyng K^+-currents in protoplasts from tobacco cell suspension cultures. J. Membrane Biol., **132**, 77-85. (1993).
Van Duijn B., Ince C., Wang Z., Weidema A. F., Libbenga K. R., Ypey D. L. Whole-cell patch-clamping; introducing substances into cells during electrical measurements from the cell membrane. A review of potential difficulties in plant and animal cells. In: 'Biotechnology application of microinjection, microscopic imaging and fluorescence', (P. H. Bach, C. H. Reynolds, J. M. Clark,P. L. Polle, J. Mottley, eds), Plenum Publishing Co., New York, USA, pp.97-108. (1993).
Van Duijn B., Ypey, D. L., Libbenga K. R. Whole-cell K^+ currents across the plasmamembrane of tobacco protoplasts from cell suspension cultures. Plant Physiol., **101**, 81-88. (1993).

The molecular analysis of higher plant embryogenesis (BIOT CT-900177)

COORDINATOR:
S.C. de VRIES, AUW, Wageningen, NL
PARTICIPANTS:
4.1.: S.C. de VRIES, AUW, Wageningen, NL
4.2.: H.A. VERHOEVEN, CPRO, Wageningen, NL
4.3.: M. TERZI, Univ. Padova, Padova, I
4.4.: K. ROBERTS, JII, Norwich, GB
4.5.: J.D. MIKKELSEN, Maribo, Kopenhagen, DK
4.6.: S.C. FRY, Univ. Edinburgh., Edinburgh, GB

BACKGROUND INFORMATION

Somatic embryogenesis is an important tool for the plant propagation industry. Formation of somatic cells that have acquired the property to form embryos is central to a better control of this process. Several lines of research suggested that somatic embryogenesis in the carrot model system involves the presence of individual secreted proteins.

OBJECTIVES AND PRIMARY APPROACHES

The central themes in this project are

a) analysis at molecular and biochemical level of several secreted glycoproteins that influence somatic embryogenesis in carrot cultures and

b) analysis of molecular markers specific for different stages of embryogenic cell formation in carrot cultures. 4.1/4.2.: Identification by automated cell tracking of single embryogenic cells, molecular, biochemical and functional analysis of 10, 32 and 47 kD secreted proteins. 4.3.: investigation of the events leading to the formation of single embryogenic cells. 4.4.: to isolate and characterise arabinogalactan-protein(s) purified from carrot suspension cell cultures, and to investigate the possible developmental role of these proteins during carrot somatic embryogenesis. 4.5.: Large scale culturing of embryogenic carrot cell cultures, purification (in mg amounts) and biochemical characterization of the 32 kD chitinase and other extracellular proteins essential for somatic embryogenesis, as well as obtaining a cDNA clone for the 32 kD chitinase. 4.6.: search for putative cell wall substrates for the 32 kD chitinase, relationship between XET activity and somatic embryogenesis.

RESULTS AND DISCUSSION

A. Groups 4.1/4.2.:

1. A secreted lipid transfer protein that marks the presence of embryogenic cells. A 10 kD secreted protein that marks the presence of embryogenic cells in carrot suspension cultures had been identified as a lipid transfer protein on the basis of cDNA derived amino acid sequence homology. Based on the amino acid sequence homology and the expression of the gene in (pro)epidermal cells, it was hypothesized that this secreted protein may be involved in cutin synthesis. The protein was purified and it was shown that the amino acid composition of the purified protein very well matched with the amino acid composition of the mature

protein as predicted from the cDNA The purified protein was shown to be able to transfer pospholipids between membrane vesicles in vitro. In addition the protein was shown to bind fatty acids and their CoA-acyl derivatives. These latter data seem to support the proposed role of this protein in the transfer of cutin monomers towards their site of polymerization.

2. Secreted proteins that influence somatic embryogenesis. Two secreted proteins, 32 kD and 47 kD, previously shown to influence the impaired somatic embryo development in the temperature sensitive carrot cell line ts11, were purified to homogeneity. The 32 kD glycoprotein consistently showed rescue activity in ts11, allowing the formation of globular and heart stage embryos at the non-permissive temperature The 32 kD protein was identified as an acidic endochitinase, based on amino acid sequence homology and catalytic properties.

The possible role of the 32 kD endochitinase in the rescue of impaired ts11 embryo deverlopment was further investigated by testing the effect of compounds consisting of or containing N-acetyl glucosamine oligomers on ts11 embryo development at the non-permissive temperature. Ts11 embryo development was shown to be stimulated by lipo-oligosaccharides from *Rhizobium*, also known as Nod-factors. The stimulatory effect was comparable with that of the 32 kD endochitinase in that both were effective only when added during the first 5 days of the embryo culture. Somatic embryogenesis in wild type carrot cell lines was not effected by the addition of the 32 kD endochitinase or the *Rhizobium* lipo-oligosaccharide. Antibodies raised against purified 32 kD endochitinase were shown to recognize several acidic proteins in suspension or embryo culture media. Some of these proteins seemed to be active chitinases, whereas others were not. The antibodies were shown to inhibit the chitinase activity of the 32 kD protein, but an initially observed inhibitory effect of the antibodies on wt carrot somatic embryogenesis was shown to be caused by an unidentified, heat-stable, low MW compound present in one particular antibody preparation.

Sequence analysis of cDNAs encoding 32 kD endochitinase revealed the presence of several highly homologous mRNAs in suspension cultured cells. The presence of at least two different 32 kD endochitinase proteins in the medium of suspension or embryo cultures was confirmed by the determined amino acid sequence of an internal peptide present in two purified 32 kD proteins. The slight difference in amino acid sequence in this peptide of the two 32 kD proteins matched with the amino acid sequences derived from different cDNA clones. Southern analysis indicated the presence of at least 4 homologous genes in carrot.

The effect of the 47 kD protein, resulting in an increased number of embryos obtained in ts11, could not be reproduced. Antibodies raised against the purified 47 kD glycoprotein were shown to recognize not only the 47 kD protein but also an ionically cell wall bound 45 kD protein. Both proteins were found present in embryogenic and non-embryogenic cell cultures. The 47 kD protein was shown to be secreted by a particular sub-class of non-embryogenic cells. Immunolocalization of the cell wall-bound 45 kD protein indicated that this protein was mainly present in cell walls between adjacent cells in non-embryogenic cell clusters. The 47 kD cDNA, cloned by screening a library with an oligonucleotide derived from the amino acid sequence of an internal peptide, was sequenced. Sequence homology with the alfalfa early nodulin gene ENOD8 was found.

3. Somatic embryos from single cells. A system to follow the fate of individual cells was developed. This involved the simultaneous testing of cell embedding techniques, media composition and the computer controlled recording of phase-

contrast and confocal scanning laser microscopic images on video tape. Using the newly developed method for tracing the developmental fate of single cells under conditions leading to the formation of somatic embryos, over 30,000 single cells were analyzed. This led to the identification of 155 single cells which formed somatic embryos. These cells could be classified into distinct morphologic classes and from each class the percentage of cells that indeed entered somatic embryogenesis were determined. It could be shown that the morphology of the intermediate proembryogenic stages differed between the different cell classes. Initial experiments to follow the developmental fate of single cells labeled with the monoclonal antibody JIM 8, previously suggested to identify a transitional cell state in the developmental pathway to somatic embryogenesis, indicated that most, if not all, cells which formed an embryo were not recognized by JIM8 at the start of the experiment. It is not yet known whether this is caused by an effect of the JIM 8 antibodies, inhibiting cells that are recognized to enter the pathway of embryogenesis, or that JIM 8-labelled cells represent a population of cells not competent to enter embryogenesis.

B. Group 4.3.:

1. The carrot cell variant ts11 cannot form embryos at 32°C but forms viable plantlets, albeit with an efficiency lower than wild type, at the permissive temperature of 24°C. It shows defects in glycosylation, transport and the formation of epidermis, which appears lax and whose cells are not closely packed. Looking at greater detail to the appearance of cells, one sees alterations in various compartments. In particular the ER is dilated and may contain electron-dense material, presumably proteins. In some cases the ER is continuous with plasma membrane and in fact, proteins, typically retained by the ER, such as BiP, are secreted by ts11 into the medium. The Golgi apparatus is much reduced but some functionality is retained because pectins and AGP are found in the medium (personal communication of S.C.Fry, University of Edinburgh); in fact secretion is much in excess compared to wild type and syntheses must have occurred in that altered Golgi. We plan to transform ts11 cells with the gene for a model glycoprotein and see where the different steps of glycosylation, glycan processing and secretion are altered. The cell wall is also altered as if some of the pectins in the matrix are secreted out and leave the cellulose fibers less compact. Alterations of the wall organization can be seen at the points of contact of the cells: in some cases a middle lamella is apparently not formed and the cells are less adherent. The picture is rather complicated and we cannot easily understand why the ts phenotype can be corrected by exogenous supply of an acidic endochitinase. We have to postulate first that the regulation of synthesis of that enzyme is altered in ts11 so that the enzyme (which is not altered per se) is synthesized, in the mutant, in lesser amounts or at later times. But then we have also to find a function (and a substrate) for an activity which is not obviously related to embryogenesis. The suggestion that the enzyme can create Nod-like signals necessary for embryogenesis is interesting and deserves further study.

2. In an attempt to define the early events in somatic embryogenesis, we explanted carrot hypocotyls in the presence of auxin and followed the modifications inducedby the phytohormone. The cells of cortical strata (epidermis, cortical parenchyma) increased in volume but not in number and eventually separated in layers from the central cylinder. Provascular cells instead extended and divided along both longitudinal and transverse planes; cells were shed off from the outer surface of the central cylinder and they constituted the progenitors of the cell line. What has been just said is true for all proliferating cell lines. Some cell lines —

because of a suboptimal induction with auxin or because they belong to a recalcitrant species or a particular genetic background — acquire only the capacity of indefinite proliferation. Other cell lines become embryogenic, which means they form pro-embryogenic masses that arise through an asymmetric division of the progenitor cells. Most interestingly, in the early stages of establishment, a pseudo-meiotic process occurs in the cell culture whereby transiently haploid cells are formed. This is demonstrated by the segregation of heterozygous markers and the resulting homozygosity of all markers in somatic embryos (obtained in collaboration with the group of V. Nuti, Istituto di Mutagenesi, CNR, Pisa). As a consequence of this random segregation of chromosomes a great variability is generated, which in part explains the phenomenon of somaclonal variation.

3. Pro-embryogenic masses, once generated in the presence of auxin, are inhibited by the continued presence of the hormone. Their physiology becomes quite different from that of the proliferating cells. They do not respond to auxin by modulating their auxin-binding proteins and the level of their plasma membrane H^+- ATPase decreases dramatically. It looks as if, by inactivating ATPase, they isolate themselves from the environment and proceed on the way of embryo formation until new structures are formed (at globular stage) that begin to react e.g. to exogenous auxin. By the time they reach plantula stage, auxin sensitivity, ATPase levels are re-established and, at the same time the level of DNA methylation, which hadbecome lower at early embryonal stages, is restored to the level characteristic of the species.

C. Group 4.4.:

Arabinogalactan-proteins (AGPs) are secreted or membrane-associated glycoproteins that have been defined as binding to -glucosyl Yariv artificial antigen, being rich in arabinose and galactose, and containing high levels of alanine, serine and hydroxyproline. Using an anti-AGP monoclonal antibody (MAC 207), bound to cyanogen bromide-activated Sepharose 4B, we have purified by immunoaffinity chromatography an extracellular AGP from the culture medium of suspension-cultured cells of *Daucus carota* (carrot). The apparent molecular weight of this highly glycosylated proteoglycan is 70-100kD as judged by SDS-polyacrylamide gels. Although its sugar analysis, -glucosyl Yariv binding, and high alanine, serine and proline content are consistent with it being an AGP, the amino acid composition analysis unexpectedly revealed this molecule to contain no hydroxyproline. This suggests that this glycoprotein is not a 'classical' AGP, but represents the first example of a new class of hydroxyproline-poor AGPs. Deglycosylation of the AGP with anhydrous hydrogen fluoride revealed that the purified proteoglycan probably contains a single core protein with an apparent molecular weight of 30kD. Direct visualisation of the native AGP in the electron microscope showed ellipsoidal putative AGP monomers, approximately 25 nm by 15nm that displayed a strong tendency to self-assemble into higher order structures. Upon dessication, the glycosylated AGP formed para-crystalline arrays, visible in the light microscope. Data from Polarised Transform Infrared microspectroscopy of these arrays suggest a more ordered structure of the carbohydrate moiety than previously proposed. Preliminary studies of the function of these proteins in relation to their possible interactions with other components of the extracellular matrix have revealed that these molecules may fulfil a role as pectin-binding proteins.

Perspectives: Despite N-terminal blockage of the AGP several internal peptide fragments have sucessfully been sequenced. AGP-specific oligonucleotides based upon someof these sequences have been made, and are currently being used to

clone the gene which encodes the AGP core protein. Once this has been achieved and the gene characterised,we shall study its pattern of expression during somatic embryogenesis.

D. Group 4.5.:

Embryogenic cell suspensions of carrot have been cultured in large scale (up to 100 l) in a rotating drum fermentor or in Erlenmeyer flasks. We have purified mg amounts of acidic chitinases from spent medium of the carrot cultures using ultrafiltration, batch and column anion-exchange chromatography, gel filtration, chromatofocusing and hydrophobic interaction chromatography. The 32 kD chitinase was purified in large scale to facilitate biochemical characterization, functional analysis of its rescue of the embryo mutant ts11 and its potential endogenous functions as well as to raise specific antibodies. By laser desorption mass spectrometry this protein was found to be 26170 ±100 D, and chromatofocusing indicated a pI of 4.2. Partial amino acid sequences were obtained from peptides generated with trypsin or endoproteinase Lys-C. These sequences cover 50% of the polypeptide chain and show significant similarities to class I and IV chitinases. They also indicate that the 32 kD chitinase contains a cysteine-rich N-terminal chitin-binding domain. This is further supported by the 32 kD chitinase being retained on a chitin column and by its high cysteine content (16 cysteines per molecule), because class II and III chitinases which lack the N-terminal chitin-binding domain have a low affinity for the chitin column and contain «8 cysteines». The antibodies raised against the 32 kD chitinase show a weak reaction with class I and IV chitinases, and vice versa, indicating that the 32 kD chitinase has some serological relationship with both class I and IV chitinases. By chromatofocusing a number of 30 kD chitinases with close serological relationship to the 32 kD chitinase were purified. Two of these proteins elute around pH 4.7 and 5.5 and are probably isoforms of the 32 kD chitinase. Furthermore, a 34 and a 29 kD chitinase have been purified from the suspension culture medium. These proteins have pIs of 4.7 and 5.0 according to chromatofocusing. Based on their MW, amino acid composition, specific chitinase activity, N-terminal sequence and serological relationship, the 29 kD chitinase probably belongs to class II, whereas the 34 kD chitinase is a class I enzyme. In order to obtain a cDNA clone for the 32 kD chitinase, PCR and subsequently 3' RACE was carried out with primers derived from the peptide sequences of the 32 kD chitinase. A large number of closely related partial clones were obtained. They can be grouped into one distantly related and four closely related types. By 5' RACE, full length clones for each of the four closely related types were obtained. These clones encode proteins which are 93-98% identical with 246 residues in the mature protein and a signal peptide of 20 or 22 residues. The proteins differ by charge i.e. the surplus of negatively charged residues is 10, 8,5 or 4, respectively, leading to calculated pIs of 4.4, 4.6, 5.1 and 5.4, respectively. The peptide sequences of the 32 kD chitinase perfectly match the most acidic of the cDNA-derived sequences.Furthermore, the calculated pI = 4.4 and MW = 26201 D correlate well with pI = 4.2 and MW = 26170 ± 100 determined for the 32 kD chitinase. Interestingly, comparison of the peptide sequences of the 32 kD chitinase and the cDNA-derived sequence indicates, that a proline located in the hinge region between the N-terminal chitin-binding domain and the catalytic domain of this chitinase has been post-translationally modified. All of the four proteins encoded by the clones show about 60% identity and 10% conserved amino acid substitutions compared to the class IV chitinase IV from sugar beet. Furthermore, they show the four deletions by which class IV chitinases differ from class I chitinases. In conclusion, the embryo-rescuing 32 kD

chitinase is encoded by a gene belonging to a gene family of at least four acidic classIV chitinase genes all of which are expressed in embryogenic carrot suspension cultures. Perspectives: The PCR clones obtained for the 32 kD and related class IV chitinases are presently used to isolate full length clones from cDNA libraries. Subsequently, transgenic plants expressing sense or antisense constructs for these class IV chitinases will provide powerful tools for the analysis of the role of these enzymes in higher plant embryogenesis.

E. Group 4.6.:

The primary plant cell wall could influence embryogenesis in two unrelated ways:
(a) a modification of wall physical properties in response to turgor pressure is likely to be critical in the determination of cell size and shape during the morphogenic events which cause embryogenesis, and
(b) the cell wall could be a source of biologically-active compounds involved in the regulation of embryogenesis.

a. ***Wall physical properties.*** A potential role for a wall-tightening enzyme (peroxidase) has been shown during early stages of embryogenesis when it is essential that cells must not be allowed to expand. However, at later stages of embryogenesis (heart-, torpedo- and cotyledon) there is the requirement for localised cell elongation and expansion, and therefore wall-loosening. During this phase we have detected an increase in the synthesis and secretion of a putative wall-loosening enzyme, xyloglucan endotransglycosylase. This cleaves a xyloglucan chain and transfers the cut end onto another (chemically similar) xyloglucan chain; it may therefore allow wall creep. By use of various carrot cell suspensions which can be manipulated to change cell shape, we have obtained evidence that this enzyme exerts its physiological activity while it passes through the load-bearing layer of the wall on its way from protoplast into the culture medium (Hetherington & Fry, Plant Physiology, In Press).

b. ***The cell wall as a source of biologically-active components.*** Work by subgroup 4.1 showed the rescue of a temperature-sensitive mutant cell line, to allow complete embryogenesis, by a 32kD chitinase from wild-type culture medium. This raises the possibility that such an enzyme may act upon the cell surface to release a biologically active molecule(s) involved in the regulation of embryogenesis.We have examined this idea by looking for putative chitinase substrates in carrot cells labelled by in vivo feeding of N-acetyl-D-[14C] glucosamine. Both cellular and medium components became radiolabelled, and acid hydrolyses followed by chromatography confirmed the label to have been incorporated essentially unaltered as hexosamine. A diverse range of such labelled preparations were then challenged with the purified 32kD chitinase. These included: polymeric fractions from culture medium, cell protein and cell wall material (obtained after phenol extraction) and living radiolabelled cells obtained from both high-density suspensions and globular embryo cultures. However, we have not been able to demonstrate any activity of the 32kD protein, which is able to attack 3H chitin, against any of the above substrates. Furthermore, we were not able to show the solubilisation of any radiolabel when the enzyme was incubated with a suspension of uniformly 14C-labelled cell walls. Gel-permeation followed by thin-layer chromatography of radiolabelled conditioned medium revealed the presence of [14C]GlcNAc-containing oligosaccharides (N-glycans) which have been shown to exert concentration-dependent physiological effects during fruit ripening.

MAJOR SCIENTIFIC BREAKTHROUGHS

4.1/4.2. Nod-factors mimic the 32 kD endochitinase in ts11 rescue. Identification of single embryogenic cells.
4.3. Identification of explant cells capable of transition into embryogenic cells.
4.4.: Structure of AGP resolved. Peptides of the core obtained.
4.5.: Identification of secreted chitinases and cloning of very closely related members of the encoding multi-gene family.
4.6.: Close correlation between XET activity and somatic embryo formation.

Significant progress has been made in most areas of the project and the objectives set have either been achieved or have exceeded expectations. The established close collaborations have been built upon and the high level of mutual interactions have resulted in numerous joint and complementary experiments.

MAJOR COOPERATIVE LINKS

General activities: Three meetings of Group 4 were held in Wageningen (May 1991), Padova (May 1992) and Norwich (May 1993), three others were held during the general meetings in Dourdan, Copenhagen and Wageningen. Specific activities: there have been annual working visits of up to three weeks of participants from group 4.1 to the laboratories of groups 4.3, 4.4 and 4.5. There have been numerous and regular exchanges of cell lines, cDNA libraries, crude and partially purified culture media, specific mono-and polyclonal antibodies, cDNA clones, purified secreted proteins, purified cell wall fragments as well as experimental protocols between all participants.

PUBLICATIONS

Joint publications

De Jong, A.J., Cordewener, J.H.C., Lo Schiavo, F., Terzi, M., Vandekerckhove, J., van Kammen, A., de Vries, S.C. (1992) A carrot somatic embryo mutant is rescued by chitinase. Plant Cell **4**, 425-433

De Jong, A.J., Heidstra, R., Spaink, H., Hartog, M.V., Hendriks, T., Lo Schiavo, F., Terzi, M., Bisseling, T., van Kammen, A., de Vries, S.C. (1993) *Rhizobium* lipooligosaccharides rescue a carrot somatic embryo mutant. Plant Cell **5**, 615-620

Pennell, R.I., Janniche, L., Scofield, G.N., Booij, H., De Vries, S.C., Roberts, K. (1992) Identification of a transitional cell state in the developmental pathway to carrot somatic embryogenesis. J. Cell Biol. **119**: 1371-1380.

Individual publications

Baldwin TC, McCann MC, Roberts K (1993) A novel hydroxyproline-deficient arabinogalactan protein secreted by suspension-cultured cells of *Daucus carota*: Purification and partial characterisation. Plant Physiology, **103**, 115-123.

Collinge, DB, Kragh, KM, Mikkelsen, JD, Nielsen, KK, Rasmussen, Vad, K. (1993) Plant chitinases. The Plant J. **3**, 31-40

De Jong, A.J., Schmidt, E.D.L., De Vries, S.C. (1993) Early events in higher plant embryogenesis. Plant Mol. Biol. **22**, 367-377

De Vries, S.C. (1992) Secreted proteins as modulators of plant embryogenesis. In: C.M. Karssen, L.C. van Loon, D. Vreugdenhil, eds., Progress in Plant Growth Regulation, pp. 340-346. Kluwer Academic Publishers, Dordrecht, The Netherlands.

De Vries, S.C., Booij, H., Cordewener, J., Van Engelen, F., Hartog, M., Hendriks, T., De Jong, A.J., Meijer, E., Sterk, P., Van Kammen, A. (1992) Cell wall glycoprotein encoding genes in somatic and zygotic embryogenesis. In: Y. Dattée, C. Dumas, A. Gallais, eds., Reproductive Biology and Plant Breeding, pp. 261-268

Filippini F, Lo Schiavo F, Terzi M, Branca C, Bellincampi D, Salvi G, Desiderio A, De Lorenzo G, Cervone F (1992) Phytoalexin elicitor-active a-1,4-D-oligogalacturonides reduce auxin perception by plant cells and tissues. In: C.M. Karssen, L.C. van Loon, D. Vreugdenhil, eds., Progress in Plant Growth Regulation, pp. 517-521. Kluwer Academic Publishers, Dordrecht, the Netherlands.
Filippini, F., Terzi, M., Cozzani, F., Lo Schiavo, F. (1992) Modulation of auxin-binding proteins in cell suspensions. II. Isolation and initial characterization of carrot cell variants impaired in somatic embryogenesis. TAG **84**, 430-434
Fry SC (1992) Xyloglucan: a metabolically dynamic polysaccharide. Trends Glycosci Glycotechnol. 4, 279-289
Fry SC, Aldington S, Hetherington PR, Aitken J (1993) Oligosaccharides as signals and substrates in the plant cell wall. Plant Physiol **103**, 1-5
Fry SC, Smith RC, Hetherington PR, Potter I (1992) Endotransglycosylation of xyloglucan: a role in cell wall yielding. Curr Topics Plant Biochem Physiol **11**, 42-62
Fry SC, Smith RC, Renwick KF, Martin DJ, Hodge SK, Matthews KJ (1992) Xyloglucan endotransglycosilase, a new cell wall loosening enzyme activity from plants. Biochem J. **282**, 821-828
Garcia-Romera I, Fry SC (1994) Absence of transglycosylation with oligogalacturonides in plant cells. Phytochemistry 35, 67-72
Hejgaard, J, Jacobsen, S, Bjorn, SE, Kragh, K (1992) Antifungal activity of chitin-binding PR4 type proteins from barley grain and stressed leaf. FEBS Lett. **307**, 389-392
Hendriks, T., Meijer, E.A., Thoma, S., Kader, J.-C., de Vries, S.C. (1994) Biochemical evidence for a role of The carrot extracellular lipid transfer protein EP2: quantitative aspects with respect to its putative role in cutin synthesis. In: Coruzzi, G., Puidgomenech, P. eds. Molecular-genetic analysis of plant metabolism and development. Plenum Publ. (in press)
Hetherington PR, Fry SC (1993) Xyloglucan endotransglycosylase activity in carrot cell suspensions during cell elongation and somatic embryogenesis. Plant Physiol. **103**, 987-992
Knox, J.P., Linstead, P.J., Peart, J., Cooper, C., Roberts, K. (1991) Developmentally regulated epitopes of cell surface arabinogalactan proteins and their relation to root tissue pattern formation. Plant J. **1**, 317-326
Lo Schiavo F, Filippini F (1992) Somatic embryogenesis as a developmental system to study modulation of auxin binding capacity. In: C.M. Karssen, L.C. van Loon, D. Vreugdenhil, eds., Progress in Plant Growth Regulation, pp. 202-206. Kluwer Academic Publishers, Dordrecht, the Netherlands.
Lo Schiavo, F., Filippini, F., Cozzani, F., Vallone, D., Terzi, M. (1991) Modulation of auxin-binding proteins in cell suspensions. I. Differential responses of carrot embryo cultures. Plant Physiol. **97**, 60-64
McQueen-Mason SJ, Fry SC, Durachko, DM, Cosgrove DJ (1993) The relationships between xyloglucan endotransglycosylase and in vitro cell wall extension in cucumber hypocotyls. Planta **190**, 327-331
Meijer, E.A., de Vries, S.C., Sterk, P., Gadella Jr., D.W.J., Wirtz, K.W.A., Hendriks, T. (1993) Characterization of the nsLTP EP2 from carrot (*Daucus carota L.*). Mol. Cell. Biochem. **123**, 159-166
Mikkelsen, JD, Berglund, L, Nielsen, KK, Christiansen, H, Bojsen, K (1992) Structure of endochitinase genes from sugar beets. In: Advances in Chitin and Chitosan. Eds. CJ Brine, PA Sandford JP Zikakis, Elsevier Appl. Sci. 344-353
Nielsen, KK, Bojsen, K, Roepstorff, P, Mikkelsen, JD (1994) A hydroxyproline-containing class IV chitinase of sugar beet is glycosylated with xylose. Plant Mol. Biol. (in press)
Nielsen, KK, Mikkelsen, JD, Kragh, KM (1994) An acidic class III chitinase in sugar beet: induction by *Cercospora beticola*, characterization and expression in transgenic tobacco plants. Mol Plant Micr. Int (in press)
Osman ME, Menzies AR, Williams PA, Phillips GO, Baldwin TC (1993) The molecular characterisation of the polysaccharide gum from *Acacia senegal*. Carbohydrate Research, **246**, 303-318.
Pennell, R.I., Janniche, L., Kjellbom, P., Scofield, G.N., Peart, J.M., Roberts, K. (1991) Developmental regulation of a plasma membrane arabinogalactan protein epitope in oilseed rape flowers. Plant Cell **3**, 1317-1326
Quesada-Allué LA, Hagelin K, Lo Schiavo F (1993) Xylose transfer to lipids by membranes of carrot cell suspensions. Phytochem. **33**, 1347-1352

Redgewell RJ, Fry SC (1993) Xyloglucan endotransglycosilase activity increases during kiwifruit (*Actinidia deliciosa*) ripening: implications for fruit softening. Plant Physiol. **103**, 1399-1406
Smallwood M, Beyen A, Donovan N, Neill SJ, Peart J, Roberts K, Knox JP (1994) Localization of cell wall proteins in relation to the developmental anatomy of the carrot root apex. Plant J. **5**, 237-246
Smith RC, Fry SC (1992) Endotransglycosylation of xyloglucan in vivo and in vitro. In: Plant cell walls as bioploymers with physiological functions. Ed. Y Masuda, pp. 313-318. Yamada Science Foundation, Osaka
Smith, R.C., Fry, S.C. (1991) Endotransglycosylation of xyloglucans in plant cell suspension cultures. Biochem. J. **279**, 529-535
Sterk, P., Booij, H., Schellekens, G.A., van Kammen, A., de Vries, S.C. (1991) Expression pattern of the carrot EP2 gene encoding an extracellular lipid transfer protein. Plant Cell **3**, 907-921
Sterk, P., De Vries, S.C. (1992) Molecular markers for plant embryos. In: SynSeeds: Application of synthetic seeds to crop improvement. CRC Press. Boca Raton, Florida, USA. pp. 115-132.
Toonen, M.A.J., Hendriks, T., Schmidt, E., Verhoeven, H., de Vries, S.C. (1994) Identification of embryo-forming single cells in carrot suspension cultures. Planta (in press)
Van Engelen, F.A., De Vries, S.C. (1992) Extracellular proteins in plant embryogenesis. Trends in Genet. **8**, 66-70
Van Engelen, F.A., De Vries, S.C. (1993) Secreted proteins in plant cell cultures. In: K.A. Roubelakis-Angelakis ed., Markers of plant morphogenesis. NATO-ASI Seminars, Plenum Publ. Co., New York. pp. 181-200.

The *rol* genes as privileged tool to study plant morphogenesis (BIOT CT-900179) (SMA)

COORDINATOR:

P. COSTANTINO, Dip. Gen. Biol. Mol., Univ. 'La Sapienza', Roma, I (DGBM)

PARTICIPANTS:

A. SPENA, Max-Planck-Institut, Köln, D (PMI)

J. TEMPÉ, Institut Sciences Végétales CNRS, Gif-sur-Yvette, F (ISV)

H. VAN ONCKELEN, Univ. Instelling Antwerpen, Antwerp, B (UIA)

D. CHRIQUI, Cytol. Embriol. Végétales. Univ. P. & M. Curie, Paris, F (CEMV)

M. BOULAY, L. Vuitton Moet Hennessy Recherche, Paris, F (LVMH)

D. MARIOTTI, Istituto Radiobiol. Ecofisiol. Vegetali CNR, Monterotondo, I (IREV)

BACKGROUND INFORMATION

Transfer of T-DNA from *Agrobacterium rhizogenes* Ri plasmids to infected plant cells causes growth of hairy roots. Among the T-DNA genes, a key role in neoplastic root differentiation and growth is played by the *rol* genes, in particular *rolA, B* and *C*. These genes are capable of triggering, to different extents, root differentiation but also affect in different and characteristic ways development and morphogenesis of transgenic plants. As morphogenesis in higher plants is still very poorly understood and 'morphogenetic' genes still quite elusive, we chose to concert efforts to exploit the unique possibility offered by a well defined set of genes capable of controlling several aspects of plant development.

OBJECTIVES AND PRIMARY APPROACHES

The objective of this project has been to exploit in a concerted fashion the *rol* gene system in shedding light on the molecular mechanisms of plant cell morphogenesis and hormonal response. The project focused on the following different aspects and scopes:

a) Regulation of the *rol* genes, in order to identify regulatory sequences and isolate plant regulatory proteins relevant to morphogenesis.

b) Biochemical role of the *rol* gene products in altering the levels of and the response to hormones, in order to shed light on some aspects of hormonal response and/or metabolism.

c) Effects of the *rol* genes on whole plants and cultured tissues, in order to better define the primary role of each *rol* gene in hairy root induction and in altering plant development.

d) Use the *rol* genes to confer to transgenic plants and tissues useful traits and/or properties (i.e. altered hormonal balance, morphology, growth pattern and rate, size, juvenility, rooting and regenerative capability etc.).

RESULTS AND DISCUSSION

The results obtained by our group as a whole have confirmed that the *rol* genes provide unique entry points on several critical and complex aspects of plant dif-

ferentiation and provide unique tools to manipulate plant growth and development.

1. Access to plant regulatory factors and genes involved in differentiation is possible through the *rol* gene system. The pattern of expression of several *rol* genes was analysed by means of gene fusions of their promoters with the GUS reporter gene. Different tissue specificity of expression of *rolA* (MPI, CEMV), *rolB* (DGBM, MPI) and *rolC* (MPI, CEMV) were pinpointed, indicating that expression of T-DNA genes, once inserted in the plant genome, is finely regulated by plant regulatory factors. In particular, an extensive cis analysis of the *rolB* promoter (DGBM) showed that this genesis activated by auxin, developmentally regulated and, in the adult plant, expressed mainly in meristematic initial cells. Several domains (denominated A, B, C, D, E) have been identified on the *rolB* promoter, different combinations of which direct selective gene expression in different populations of cells of the root meristem (DGBM). One of these domains (domain B) is necessary for expression in all meristematic cell types. Gel mobility shift analysis and screening of an untransformed tobacco expression library with a tetramer of domain B led to the isolation of a cDNA encoding for a plant regulatory protein (denominated *rBBf*) specifically binding to domain B. This regulatory factor, currently being characterized, must regulate other plant genes which are presumably involved in morphogenesis and whose identification will be possible via further work on *rBBf* (DGBM). By using affinity-purified antibodies the cell-specificity of the *rolC* promoter has been investigated (MPI, CEMV). During vegetative growth, *rolC* is expessed only in companion cells and, in roots, in the initials of protophloem cells. It is important to stress that under no circumstances the use of GUS staining was giving a precise cellular localization of promoter activity. The same study has shown that the sites of action of the *rolC* gene are distinct from the site of expression, results compatible with models involving transported signal molecules in the mode of action of *rolC* (MPI, CEMV). Although not expressing *rolC*, pith and leaf parenchyma cells of *rolC* transgenic plants were in fact affected in their growth parameters, indicating correlative information transported from the site of expression to the site of action (MPI, CEMV).

2. Investigation on the biochemical role of the *rol* gene products has demonstrated that plant physiology and development can be modified by enzymatic systems capable to interfere with hormone biosynthesis and/or release from inactive conjugates. In order to understand the function of *rolA, B* and *C*, the three genes have been expressed in plants, bacteria and yeast (MPI). Expression in transgenic tobacco plants has been performed under the control of their own promoters, under the control of a constitutive promoter (i.e. 35S from Cauliflower mosaic virus) and under the control of a tissue-specific promoter (i.e. tapetum-specific promoter). Expression in bacteria has been achieved by using an inducible promoter (MPI, DGBM). Expression in yeast has been achieved by using an inducible promoter (MPI). To facilitate the study of the function and expression of *rolA, B* and *C*, antibodies against all three gene products have been raised in rabbits and affinity-purified (MPI, DGBM). These biochemical and immunological approaches allowed to conclude that the *rolB* peptide is a cytosolic enzyme able to hydrolyse *in vitro* indoxyl-glucoside and other glucosides (e.g. cytokinin O-glucosides; indoxyl- galactoside) (MPI). The *rolC* peptide is a cytosolic enzyme able to hydrolyse *in vitro* cytokinin N- and O-glucosides (MPI). The *rolA* gene has been mutagenised (EMS) and null mutations pinpointing amino acids relevant for its function have been identified (MPI). On the other hand, work at DGBM and ISV had shown that *rolB* enances by several orders of magnitude the sensitivity to auxin of transformed cells. Accordingly, subsequent work at DGBM showed that

mebrane preparations from *rolB*-transformed cells bind substantially more auxin than normal ones, an evidence of a possible role of *rolB* in altering the auxin reception mechanism. A major part of the work at UIA involved the elaboration of sophisticated analytical techniques (solid phase extraction, preparative and analytical HPLC with on line fluorometry and diode array, HPLC-MS with thermo/ plasma spray interfaces, GC-MS and raising polyclonal antibodies for immune affinity purification and RIA purposes) for the qualitative and quantitative analysis of plant growth regulators and their metabolites. These techniques have been used to study the functional relationship between altered endogenous hormone concentrations and metabolism and physiological and morphological traits specifically observed in plants transformed with Ri T-DNA genes which affect hormone metabolism and/or sensitivity. The 'hairy-root' specific morphological traits of Ri T-DNA- transgenic tobacco plants were specifically correlated with an alteration of hormonal levels in the stem apices due to: drastic accumulation of cytokinins at a very young age, attenuation of the basipetal auxin gradient and absence of the transient ABA accumulation shown in the apices of untransformed plants during the vegetative developmental stage. The work at ISV has shown that the product encoded by Ri ORF13 modifies plant development by modification of cellular responses to auxin and cytokinin. As the morphological alterations due to overexpression of ORF13 in transgenic plants are transmittable through a graft junction, it is likely that ORF13 is responsible for the synthesis of an hitherto unknown, transportable morphogenic substance. As plant cells recognize it, this product is either a normal metabolite present in abnormally high concentrations, or the analogue of a normal plant product involved in hormone response. A bioassay is currently being developed for purifying the unknown ORF13 product.

3. Basic aspects of plant physiology could be addressed through the analysis of the effects of the *rol* genes in transgenic plants and cultured tissues. The characteristic morphological aberrations produced by expression and overexpression of the *rol* and other RiT-DNA genes have been revisited and analyzed in detail in order to clarify the primary role of these genes in controlling plant development. The effects of *rolB* on *in vitro* morphogenesis have been analyzed at DGBM by means of thin cell layers (TCLs). TCLs from *rolB*-transformed and untransformed tobacco plants have been cultured under conditions of root, flower and shoot neoformation. Unexpectedly, *rolB* was shown to enhance not only the rooting programme but also, to an equivalent extent, flower and shoot neoformation. In all cases the number of neoformed meristems in *rolB*-TCL sis greatly increased as compared to untransformed controls. Thus, in contrast to what has been held so far, *rolB* should not be regarded as a 'root inducing' gene but rather as a gene capable of promoting meristem formation. The differentiative fate of these meristems will then depend upon the prevailing hormonal balance and/or cell commitment. Further work is now in progress to exploit the *rolB*-enhanced meristem formation in order to get access to plant genes involved in the genesis and organization of meristems (DGBM). Cellular and morphogenetical approaches were developed at CEMV and MPI on *rolA* and *rolC* transgenic tobaccos at various developmental stages in order to understand their modified phenotypes and to correlate the observed effects with the *rol* gene functions and sites of expression. The tissue specificity of expression and its developmental regulation were also examined using transgenic tobacco plants harbouring *rolA* and *rolC*-GUS transcriptional fusions in order to establish relationships between the changes at cell and tissue levels of organization and the sites of gene expression. Plants expressing *rolA* or GUS linked to various functional domains of the *rolA* promoter were also examined respectively for their developmental patterns and for the sites

of expression directed by each domain. Most of the histological findings on *rolC*-transgenic plants were found to be reminiscent of a cytokinin effect, as expected from the function of cytokinin-glucosidase of the *rolC* gene product (see above).

4. The *rol* genes have also been utilized as morphogenes for applicative purposes. At LVMH the possible commercial exploitation of the morphological alterations induced by *rolB* on roses were investigated. The hybrid G. Delbard registered cultivar is used for cut flower rose production. One agronomical characteristic of interest for such genotype is productivity, which is in part correlated with stem number. *rolC* is known to induce morphological changes such as lateral branching, shortened internodes and precocious flowering. Thus *rolC* has been introduced in GD roses under the control of its own regulatory sequence, in collaboration with MPI. Among 400 transformed plants, 50 were cultivated in growth chamber and their agronomical characteristics were monitored. Some of them had abnormal phenotypes: a few showed a significative increase in stem number (2 to 3-fold as compared to the controls) and were considered interesting in terms of productivity. Currently, a vast greenhouse trial is being performed on 420 transgenic roses. Parameters being evaluated are: plant shape, stem length, stem number petal number, flower colour and time course of flowering (LVMH). At IREV, the potentialities of *rolB* in enhancing propagation of woody species via cuttings was explored. *Actinidia deliciosa* (kiwi) was used as a model species. Cuttings of *rolB*-transgenic plants were shown to produce significantly more roots than untransformed controls. Rooting was anticipated and significantly higher than in controls were the number of rooted cuttings as well as the number of roots per cutting. Chestnut (*Castanea sativa*) cuttings from valuable centuries-old plants were locally treated with suspension cultures of *Agrobacterium* strains carrying *rolB* and abundant rooting of the cuttings was obtained. It must be pointed out that this is the first successful attempt in adult chestnut cutting propagation (IREV). Finally, *Medicago sativa* (alfalfa) transgenic for *rolB* showed a more developed root system and a doubled number of stems as compared to controls (IREV), confirming the potentialities of the *rol* genes in breeding schemes of agronomically relevant plants.

MAJOR SCIENTIFIC BREAKTHROUGHS

First identification of a plant regulatory gene that confers gene expression in meristems (DGBM). Demonstration that the primary effect of *rolB* is in promoting meristem formation (DGBM). Definition of the enzymatic activities of the *rolB* (MPI + UIA) and *rolC* proteins (MPI + UIA). Localization of cell-specific expression of *rolA* (MPI + CEMV), *rolB* (DGBM, MPI) and *rolC* (MPI + CEMV). Correlation of hormone levels with *rol* genes-induced morphological and physiological modifications in transgenic plants (MPI + UIA). Identification of possibly a novel plant growth factor as the product of Ri ORF13 (ISV + UIA). *rolB* was shown to significantly enhance rooting in cuttings of agronomically relevant recalcitrant woody species (kiwi, chestnut) (IREV + DGBM). *rolC* has been used in modifying growth flowering of roses (LVMH + MPI).

MAJOR COOPERATIVE LINKS

Most projects have been carried out in collaboration between different labs (see above). Cooperative links were also maintained by free access to all tools, by two annual meetings of all the group leaders and by several bilateral personnel exchanges and visits.

PUBLICATIONS

Joint publications

Spena, A., Prinsen, E., Fladung, M., Schulze, S.C. and Van Onckelen, H.: The indoleaceticlysine synthetase gene of *Pseudomonas syringae* subsp. *savastanoi* induces developmental alterations in transgenic tobacco and potato plants. Mol. Gen. Genet. **227**, 205-212, 1991.

Estruch, J.J., Chriqui, D., Grossmann, K., Schell, J. and Spena, A.: The plant oncogene *rolC* is responsible for the release of cytokinin from glucoside conjugates. The EMBO Journal **10**, 2889-2995, 1991.

Estruch, J.J., Prinsen, E., Van Onckelen, H., Schell, J. and Spena, A.: *Viviparous* leaves produced by somatic activation of an inactive cytokinin-synthesizing gene. SCIENCE **254**, 1364-1367, 1991.

Maurel, C., Barbier-Brygoo, H., Spena, A., Tempé, J. and Guern, J.: Single *rol* genes from the *Agrobacterium rhizogenes* TL-DNA modify the response to auxin of tobacco mesophyll protoplasts. Plant Physiol. **97**, 212-216, 1991.

Prinsen, E., Chauvaux, N., Schmidt, J., John, M., Wieneke, U., De Greef, J., Schell, J. and Van Onckelen, H.: Stimulation of indole-B-aceticacid production in *Rhizobium* by flavonoids. FEBS Lett. **1**, 53-55, 1991.

Spena, A., Estruch, J.J., Aalen, R.D., Prinsen, E., Parets-Soler, A., Nacken, W., Sommer, H., Chriqui, D., Grossmann, K., Van Onckelen, H. and Schell, J.: Transgenic plants and transgenic plant mosaics for the expression of pathogen derived genes able to affect phytohormone activity. In: Progress in Plant Growth Regulation, Eds. C.M. Karsen, L.C. Van Loon and D. Vreugdenhil, Kluwer Acad. Publ.1992, pp. 724- 730.

Prinsen, E., Berchetche, J., Chriqui, D. and Van Onckelen, H.: *Pisum sativum* epicotyls inoculated with *Agrobacterium rhizogenes* agropine strains harbouring various T-DNA fragments: morphology, histology and endogenous indole-3-acetic acid and indole-3-acetamide content. J. Plant Physiol. **140**, 75-83, 1992.

Spena, A., Estruch, J.J., Nacken, W., Prinsen, E., Van Onckelen, H. and Sommer, H.: Anther-specific expression of the *rolB* gene of *Agrobacterium rhizogenes* increases IAA content in anthers and alters anther development. T.A.G. **84**, 520-527, 1992.

Mengoli, M., Chriqui, D. and Bagni, N.: Protein,free aminoacid and polyamine content during development of hairy root *Nicotiana tabacum* plants. J. Plant Physiol. **139**, 697-702, 1992.

Mengoli, M., Ghelli, A., Chriqui, D. and Bagni, N.: Growth kinetics, polyamine pattern and biosynthesis in hairy root lines of *Nicotiana tabacum* plants. Physiol. Plant. **85**, 697-703, 1992.

Mengoli,M., Chriqui, D. and Bagni, N.: Putrescine biosynthesis and oxydation in normal and hairy root tobacco plants. J. Plant Physiol. **140**, 153-155, 1992.

Spena, A., Estruch, J.J., Nacken, W., Prinsen, E., Van Onckelen, H. and Sommer, H.: Anther-specific expression of the *rolB* gene of *Agrobacterium rhizogenes* increses IAA content in anthers and alters anther development. T.A.G. (in press).

Chriqui, D., Guivarc'h, A., Caissard, J.C., Dewitte, W. and Van Onckelen, H.: Cellular parameters involved in host susceptibility to various agrobacteria. Giorn. Bot. Ital. **127**, 255-263, 1993.

Estruch, J.J., Granell, A., Hansen, G., Prinsen, E., Redig, P., Van Onckelen, H., Schwarz-Sommer, Z., Sommer, H. and Spena, A.: Floral development and expression of floral homeotic genes are influenced by cytokinins. Plant J., **4**, 379-384, 1993.

Guivarc'h, A., Caissard, J.C., Brown, S., Marie, D., Dewitte, W., Van Onckelen, H. and Chriqui, D.: Localizationoftargetcells and improvement of *Agrobacterium*-mediated transformation efficiency by direct acetosyringone pretreatment of carrot discs. Protoplasma, **174**, 10-18, 1993.

Tepfer, M., Vilaine, F., Carneiro, M., Pautot, V., Rossignol, M., Van Onckelen, H., Legué, V., Perbal, G., Besnard, C., Noin, M., Guivarc'h, A. and Chriqui, D.: Physiological and morphological consequences of expression of the *rolA* gene in transgenic tobacco plants. Actual. Bot. Gallica, **140**, 685-691, 1993.

Individual publications

Beinsberger, S., Valcke, R., Deblaere, R., Clijsters, H., De Greef, J. and Van Onckelen, H.: Effects of the introduction of *Agrobacterium tumefaciens* T-DNA *ipt* gene in *Nicotiana tabacum* cv Petite havana SR1 plant cells. Plant Cell Physiol. **32**, 489-496, 1991.

Heylen C., Vendrig J. and Van Onckelen H.: The accumulation and metabolism of plant growth regulators during organogenesis of thin cell layers of *Nicotiana tabacum*. Physiol. Plant. **83**, 578-584, 1991.

Rugini, E., Pellegrineschi, A., Mencuccini, M. and Mariotti, D.: Increase of rooting ability in the woody species of kiwi (*Actnidia deliciosa* A Chev.) by transformation with *Agrobacterium rhizogenes rol* genes. Plant Cell Rep. **10**, 291-295, 1991.
Estruch, J.J., Parets-Soler, A., Schmülling, T. and Spena, A.: Cytosolic localization in transgenic plants of the *rolC* peptide from *Agrobacterium rhizogenes*. Plant Mol. Biol. **17**, 547-550, 1991
Estruch, J.J., Schell, J. and Spena, A.: The protein encoded by the *rolB* plant oncogene hydrolyses indole glucosides. The EMBO Journal **10**, 3125-3128, 1991
Hansen, G., Larribe, M., Vaubert, D., Tempé, J., Bierma, B.J., Montoya, A.L., Chilton, M.D. and Brevet, J.: *Agrobacterium rhizogenes* pRi8196T-DNA: mapping and DNA sequence of functions involved in mannopine synthesis and hairy root differentiation. Proc. Natl. Acad. Sci. U.S.A. **88**, 7763-7767, 1991.
Capone, I., Cardarelli, M., Mariotti, D., Pomponi, M., De Paolis, A. and Costantino, P.: Different promoter regions control level and tissue specificity of expression of *Agrobacterium rhizogenes rolB*gene in plants. Plant Mol. Biol. **16**, 427-436, 1991.
Altamura, M.M., Archilletti, T., Capone, I. and Costantino, P. Histological analysis of the expression of *Agrobacterium rhizogenes rolB*-GUS gene fusions in transgenic tobacco. New Phytol. **118**, 69-78, 1991.
Costantino, P., Cardarelli, M., Capone, I., De Paolis, A., Filetici, P., Pomponi, M. and Trovato, M. Morphogenetic genes in the T-DNA of Ri plasmids, in: Plant Molecular Biology 2, Herrmann, R.G.and Larkins, B., eds. Plenum Press, New York, pp. 211-218, 1991.
Spena, A. and Schulze, S.C.: Chimaeras and transgenic plant mosaics: a new tool inplant biology. In: Advances in Molecular Genetics of Plant-Microbe Interactions. Hennecke, H. and Verma, D.P.S. eds., Kluwer Academic Publ. **1**, 352-355, 1991.
Dessaux, Y., Petit, A. and Tempé, J.: Opines in the biology of *Agrobacterium*. In: Molecular Signals in Plant-Microbe Communications, pp. 109-136. Verma, D.P.S. ed., CRC Press, Boca Raton, 1991
Phelep, M., Petit, A., Martin, L., Duyhoux, E. and Tempé, J.: Transformation and regeneration of a nitrogen-fixing tree *Allocasuarina verticillata* Lam. Biotechnology **9**, 461-466, 1991.
Van der Linden, A., Gadeyne, J., Van Onckelen, H., Van Laere, A. and Decleir, W.: Involvement of cyclic nucleotides inlight-induced resumption of development of *Artemia* embryos. J. Exp. Zool. **258**, 312-321, 1991
Marcussen, J., Seiden, P., Olsen, C. and Van Onckelen, H.: Monoclonal antibodies toward a stable cyclic AMP-C18 antigen. Anal. Biochem. 198, 324-329, 1991.
Mariotti, D., Fontana, G.S., Mencuccini, M., Caretto, S., Rugini, E., Nicolodi, C. and Santini, L.: *Agrobacterium rhizogenes rol* genes: an alternative approach to improve root formation in micropropagated woody species. Giorn. Bot. Ital., **125**, 859-864, 1991.
Rugini, E. and Mariotti, D.: *Agrobacterium rhizogenes* T-DNA genes and rooting in woody species. Acta Hort., **300**, 301-308, 1991.
Beinsberger, S., Valcke, R., Clijsters, H., De Greef, J. and Van Onckelen, H.: Effects of enhanced cytokinin levels in *ipt* transgenic tobacco. In : Physiology and Biochemistry of Cytokinins in Plants, Eds. M. Kaminek, D.W.S. Mok and E. Zazimalova, SPB Publ. Comp., pp. 77-82, 1992.
Beinsberger, S., Valcke, R., Clijsters, H., De Greef, J. and Van Onckelen, H.: Morphological characteristics and phytohormone content of *ipt*-transgenic tobacco. In: Progress in Plant Growth Regulation, Eds. C.M. Karsen, L.C. Van Loon and D. Vreugdenhil, Kluwer Acad. Publ.1992, pp. 738-745.
Costacurta, A., Prinsen, E., Van Onckelen, H., Michiels, K., Vanderleyden, J. and Nuti, M.P.: IAA synthesis in *Azospirillum brasiliense* Sp6: analysis of a mutant impaired in IAM-hydrolase. Symbiosis, **13**, 151-158, 1992.
Caissard, J.C., Rembur, J. and Chriqui, D. : Electron microscopy and X-ray microanalysis as tools for fine localization of the β-glucuronidase activity in transgenic plants harbouring the GUS reporter gene. Protoplasma, **170**, 68-76, 1992
Chichiriccò, G., Costantino, P. and Spanò, L.: Expression of the *rolB* oncogene from *Agrobacterium rhizogenes* during zygotic embryogenesis in tobacco. Plant Cell Physiol. **33**, 827-832, 1992.
Spena, A., Estruch, J.J. and Schell, J.: On microbes and plants: new insights in phytohormonal research. Curr. Opin. in Biotech., **3**, 159-163, 1992

Spena, A., Estruch, J.J., Hansen, G., Lagenkemper, K., Berger, S. and Schell, J.: The *rhizogenes* tale: modification of plant growth and physiology by an enzymatic system of hydrolysis of phytohormone conjugates. In: Advances in Molecular Genetics of Plant-Microbe Interactions, Nester, E.W. and Verma, D.P.S. eds., Kluwer Acad. Publishers, pp. 109-124, 1992
Hansen, G., Tempé, J. and Brevet, J.: A T-DNA transfer stimulator sequence in the vicinity oftheright border of pRi8196. Plant Mol. Biol. **20**, 113-122, 1992
Dessaux,Y., Petit, A. and Tempé, J.: Opines in *Agrobacterium* biology. In: Molecular Signals in Plant-Microbe Communications, Verma, D.P.S. ed., pp. 109-136, 1992
David, C. and Tempé, J.: Transformation of *Catharantus*. In: Biotechnology in Agriculture and Forestry. Bajaj, Y.P.S. ed. Springer Verlag, Berlin, 1992.
Lopez-Carbonell, M., Alegre, L., Prinsen, E. and Van Onckelen, H.: Diurnal fluctuations of endogenous IAA content in *Aralia* leaves. Biol. Plant., **34**, 223-227, 1992.
De Agazio, M., Decesare, F. and Mariotti, D.: Polyamine content and rhizogenesis in tobacco transformed by the *rolB* gene from *Agrobacterium rhizogenes* T-DNA. Life Sci. Adv.-Plant Physiol., **11**, 5-11, 1992.
Mariotti, D., Santini, L. and Spanò, L.: Rooting in *Agrobacterium rhizogenes*-transformed tobacco tissues: morphogenetic responses related to auxin metabolism. Life Sci. Adv.-Plant Physiol.,**11**,13-18, 1992.
Guyon,P.,Petit,A.,Tempé,J. and Dessaux, Y.: Tranformed plants producing opines specifically promote growth of opine-degrading *Agrobacteria*. Plant-Microbe Interact. **6**, 92-98, 1993
Chauvaux, N., Van Dongen, W., Esmans, E. and Van Onckelen, H.: Liquid chromatographic mass spectrometric determination of 1-aminocyclopropane-1-carboxylic acid in tobacco. J. Chromat., **647**, 211-218, 1993.
Prinsen, E., Costacurta, A., Michiels, K., Vanderleyden, H. and Van Onckelen, H.: *Azospirillum brasiliense* indole-3-acetic acid biosynthesis: evidence for non-tryptophan dependent pathway. Mol. Plant-Microbe Interact. **6**, 601-609, 1993.
Van Loven, Beinsberger, S., Valcke, R., Van Onckelen, H. and Clijsters, H.: Morphometric analysis of the growth of phsp70-ipt transgenic plants. J. Exp. Bot., **44**, 1671-1678, 1993.
Zhou, X., Costantino, P. and De Paolis, A.: Isolation and sequence analysis of cDNA clones encoding the precursor of the 10-kDa polypeptide of photosystem II from tobacco. J. Genet. & Breed., **47**, 273-276, 1993.
Costantino, P., Capone, I., Cardarelli, M., De Paolis, A. and Trovato, M. *rolB*: a bacterial gene capable of controlling auxin response and morphogenesis in plant cells. In: Biochemical Mechanisms Involved in Plant Growth Regulation, Smith, C.J. ed., Oxford University Press, pp. 171-177, 1993.
Rinallo, C. and Mariotti, D.: Rooting in *Castanea sativa* shoots: effect of *Agrobacteriumrhizogenes* T-DNA genes. J. Hort. Sci., **66**, 16-23, 1993.
Rinallo, C. and Mariotti, D.: *Agrobacterium* as a tool for gene transfer in woody species: the state ofthe art and practical perspectives. Giorn. Bot. Ital., **128**, 1-12, 1993.
Haagale, T., Santini, L. and Mariotti, D.: *Agrobacterium*-mediated transformation in trees: preliminary studies on the transfer of *rol* genes into some north European woody species. Adv. Hort. Sci., **7**, 1-4, 1993.
Altamura, M.M., Capitani, F., Gazza, L., Capone, I. and Costantino, P.: The plant oncogene *rolB* stimulates the formation of flower and root meristemoids in tobacco thin cell layers. New Phytol., **126**, 283-293, 1994.
Costantino, P., Capone, I., Cardarelli, M., De Paolis, A. and Trovato, M. The *rol* plant oncogenes: basic and applied aspects. In: Applications of DNA technology to Agriculture and Health, Riazuddin, S. and Papamatheakis J. eds., University of the Punjab Press, pp. 77-87, 1994.
Caissard, J.C., Guivarc'h, A., Rembur, J. and Chriqui, D.: Spurious localization of diX-indigo microcrystals generated by the histochemical GUS assay. Trans. Res., **3**, 176-181, 1994.
Filippini, F., Lo Schiavo, F., Terzi, M., Costantino, P. and Trovato, M.: The plant oncogene *rolB*alters binding of auxin to plant cell membranes. Plant Cell Physiol., **35**, 93-98, 1994.
Carbonell, M., Alegre, L. and Van Onckelen, H.: Effects of water stress on cellular ultrastructure and on concentrations of endogenous abscisic acid and indole-3-aceticacidin *Fatsia-japonica* leaves. Plant Growth Reg., **14**, 29-35, 1994.
Capone, I., Frugis, G., Costantino, P. and Cardarelli, M.: Expression in different populations of cells of the root meristem is controlled by different domains of the *rolB* promoter. Plant Mol. Biol., (in press).

Prinsen, E., Chriqui, D., Vilaine, F., Tepfer, M. and Van Onckelen, H.: Endogenous Phytohormones in tobacco plants transformed with *Agrobacterium rhizogenes* TL-DNA. J. Plant Physiol., (in press).
Serino, G., Clerot, D., Brevet, J., Costantino, P. and Cardarelli, M.: *rol* genes of *Agrobacterium rhizogenes* cucumopine strain: sequence, effects and pattern of expression. Plant Mol. Biol., (in press).
Costantino, P., Capone, I., Cardarelli, M., De Paolis, A., Mauro, M.L. and Trovato, M.: Bacterial plant oncogenes: the *rol* genes' saga. Genetica, (in press).
Altamura, M.M., Capitani, F., Tomassi, M., Capone, I. and Costantino, P.: The plant oncogene *rolB* enhances meristem formation in tobacco thin cell layers. (in press).
Prinsen, E., Redig, P., Galis, I., Van Dongen, W. and van Onckelen, H.: Quantifying phytohormones in transformed plants. In: Methods in Molecular Biology, vol. *Agrobacterium* protocols. Eds. K. Gartland and M. Davey, Humana Press Inc., New Jersey, In Press
Spena, A., Estruch, J.J. and Schell, J.: On microbes and plants: new insights in phytohormonal research. Current Opinion in Biotechnology. (in press).
Tempé J.: Introduction of foreign genetic information in plants. Acta Hort. (in press).
Frugis, G., Caretto, S., Santini, L.andMariotti, D.: *Agrobacterium rhizogenesrol* genes induce productivity-related phenotypical modificationsin 'creeping-rooted' alfalfa types. Plant Cell Rep., (in press).

Molecular analysis of auxin-specific signal transduction in plant cell communication (BIOT CT-9000178)

COORDINATOR:
K. PALME, Max-Planck-Institut, Köln, D (MPI)
PARTICIPANTS:
M. VENIS, Horticulture Research Int., Kent, GB (HRI)
H. BARBIER-BRYGOO, Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, F (ISV)
D. INZÉ, Rijksuniversiteit Gent, Gent, B (RUG)
P.A. MILLNER, University of Leeds, Leeds, GB (ULeeds)
F. OKKELS, Danisco/AS, Copenhagen, DK (Danisco)

BACKGROUND INFORMATION

Auxin is a phytohormone that is involved in the control of numerous aspects of growth and development in higher plants, including cell division, stem elongation, xylem differentiation, gravitropism and senescence. This phytohormone is of prime importance for regeneration of plants from single cells. Receptor-like proteins have been assumed to play an important role in recognition and transmission of the auxin signal. Aiming to isolate and characterize such receptors or other elements of the auxin signalling chain is of general relevance for providing the basic information to making in the long run plant regeneration a more predictable process.

OBJECTIVES AND PRIMARY APPROACHES

Molecular, structural and functional analysis of ER located auxin binding proteins (MPI). Use of auxin-specific photoaffinity labeling to identify novel proteins involved hormone action (MPI, Danisco). Production of antibodies to auxin binding site or other important epitopes. Epitope mapping (HRI). ABP interactions with cellular partners, analysis of effectors in functional responses (HRI). Functional study of abp1-related proteins at the plasma membrane (ISV). Identification of the molecular mechanisms involved in the electrical response to auxin (ISV). Molecular characterization of components of auxin perception units at the plasma membrane (ISV). Uncoupling of cell divison and morphogenesis (RUG). Expression of cell cycle genes in *Arabidopsis thaliana* (RUG). GTP binding proteins, receptors and nucleotide diphosphate kinase activity associated with higher plant membranes (ULeeds). Use of synthetic oligopeptides corresponding to the auxin binding protein C-terminus as agonists for the auxin signalling cascade (ULeeds).

RESULTS AND DISCUSSION

A. MPI:

We have analyzed auxin-binding proteins from maize encoded by the *ZmERabp* gene family. Open reading frames of cDNA clones predict proteins containing N-terminal hydrophobic signal sequences. *In vitro* studies show that the ZmERabp1 protein can be translocated into ER-derived microsomes where it is processed and glycosylated. A cDNA clone encoding the ZmERabp4 protein predicts an open reading frame with a signal sequence that shows striking differences in charge distribution, in comparison to the signal sequence of ZmERabp1. Two translation products are synthesized from the ZmERabp4 transcript in the *in vitro* system, but

only one of them is translocated into maize endosperm microsomes, indicating that specific cotranslational modifications in the primary sequence remaining after processing may play a role in the cellular trafficking of the ZmERabp4 protein.

Three maize genes encoding 22 kDa auxin binding proteins (abp) were isolated and sequenced. One gene corresponds to the abp1 gene (*ZmERabp1*). The two other genes *ZmERabp4* and *ZmERabp5* are closely related to each other, but their 5' flanks are very different. *ZmERabp5* corresponds to a recently published partial cDNA. RNA gel blot analysis using gene-specific probes indicates that ZmERabp4 is expressed in leaves and coleoptiles but weakly in roots. ZmERabp5 is slightly expressed in these tissues. Primer extension analysis reveals that the *ZmERabp1* leader is 320 bases. Transient promoter- reporter gene expression in maize leaf protoplasts indicates that the relative expression of *ZmERabp1*, *ZmERabp4* and *ZmERabp5* is 6.4: 1.8: 1.0. Promoter deletion analysis of *ZmERabp1* has identified a region from -684 and -450 (relative to ATG) which when deleted results in two-fold higher expression. This region contains sequences related to both an enhancer and a G-box. Another region between -449 and -258, which contains the TATA box and transcription start, results in a large decrease in expression when deleted.

A *Zea mays* cDNA clone, *ZmERabp4*, was further studied. The primary amino acid sequence contains an N-terminal hydrophobic leader sequence, a potential glycosylation site (Asn^{136}- Thr-Thr) and a C-terminal KDEL motif known to be responsible for retention of proteins within the lumen of the ER. The expression pattern of the *ZmERabp4* gene in various organs of maize differs from the expression pattern previously observed for the *ZmERabp1* gene. The *ZmERabp4* gene is expressed highly in male flower organs, whereas the *ZmERabp1* gene shows highest expression in female flower parts. *In situ* hybridisation and analysis by laser scanning microscopy revealed enhanced levels of expression for both genes in the coleoptile when compared with the primary leaf of etiolated maize seedlings.

B. Danisco, MPI:
We used a biologically active and photolabile auxin analogue, 5'-azido-[7- ^{3}H]indole-3-acetic acid ([^{3}H]N_3IAA), to search for auxin-binding proteins in cytosolic extracts from maize coleoptiles and identified a protein with a molecular mass of 60 kDa (p60). Binding of [^{3}H]N_3IAA is highly specific as demonstrated by competition analysis with functionally relevant auxin analogues. p60 is found in coleoptiles and roots of etiolated maize seedlings and was detected in cytosolic as well as in microsomal fractions. The protein binds to 1-naphthylacetic acid (1-NAA) sepharose and is eluted with auxins. A purification scheme resulting in homogenous p60 protein was devised and it was shown that p60 has β-D- glucoside glucohydrolase activity (E.C. 3.2.1.21). The hydrolytic activity of p60 for the synthetic substrate p-nitro-phenyl-β-D-glucopyranoside is diminished by 1-NAA. p60 shows high substrate specificity, it hydrolyzes indoxyl-O-glucoside, but not β-(1,4)-cellobiose, IAA-inositol or IAA-amino acid conjugates. Inactive phytohormone conjugates are abundant in plant tissues, but their normal biological functions remain obscure. The apparent physiological activity of any particular conjugate correlates with its rate of hydrolysis in plant tissues. Analysis of the bacterial pathogen *Agrobacterium rhizogenes* has shown that abnormal development in plants may be promoted through the action of the *rolC* oncogene, which encodes a β-glucosidase that can release free cytokinin from inactive conjugates. Despite this demonstration that phytohormone conjugates can be exploited to modify plant development, we lack knowledge about hydrolases in the normal plant that are capable of releasing phytohormones. p60 was purified, partially sequenced and its cDNA, termed *Zmp60.1*, isolated. The protein encoded by this cDNA, Zm-p60.1,

belongs to such a family of hydrolases. The significane of similarities of peptide similarities between Zmp60 peptides and *rolC* was intriguing, but full comparison of structures did not reveal further structural similarities. However, despite distinct phylogenies, they seem to have similarities in their enzymatic properties. Recombinant Zmp60.1 expressed in *E. coli* showed no or low activities on substrates from several groups of naturally occurring glycosides including disaccharides involved in plant cell wall degradation (cellobiose, laminiaribiose), phenolic glucosides (salicin) and flavonoid glycosides (rutin). Prompted by the presence of Zmp60.1 in young maize seedlings and by its original identification through photoaffinity labeling with an azido derivative of IAA, we tested whether commmon phytohormone conjugates could be hydrolyzed. We were not able to demonstrate hydrolysis of IA-glucoseester by Zmp60.1, consistent with its relative specificity for the glycosidic bond. However, cytokinin-O- and -N3-glucosides, but not cytokinin-N7 and -N9-glucosides, were hydrolyzed by both recombinant and plant p60. In plant tissues, O-glucosides are the major metabolizable conjugates form of cytokinin from which cytokinins can be released by endogenous hydrolases. The interesting observation that the *rolC* protein, but not Zmp60.1 is able to hydrolyze zeatin-N7- and -N9-glucosides can be rationalized with regard to the physiological roles proposed for each enzyme; cytokinin glucosylation on N7 and N9 is considered to be one of the mechanisms for irreversible cytokinin inactivation in plant cells, and thus Zmp60.1, an endogenous plant enzyme, does not attack substances which accumulate as inactivated end products under normal physiological conditions. In contrast, the *rolC* protein is part of the mechanism by which the plant pathogen *Agrobacterium rhizogenes* subverts normal plant development to its own advantage; there is no obvious reason for the activity of such an enzyme to be restricted to those

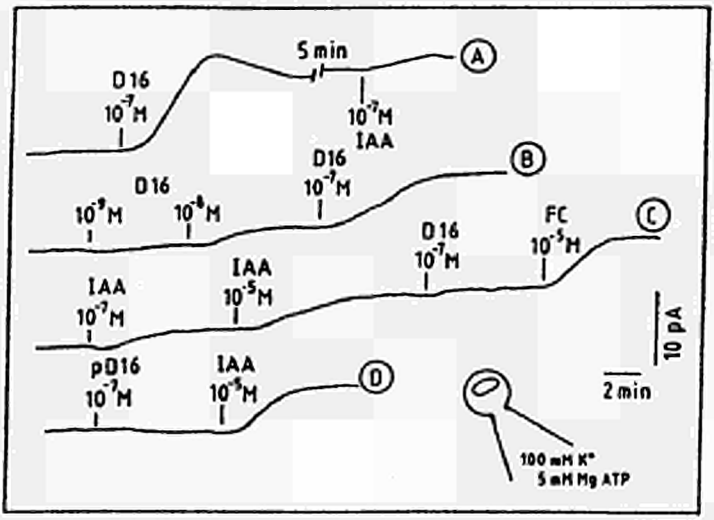


Fig. 1. Effect of antipeptide antibodies on the electrical membrane current recorded from Zea mays *protoplasts with the patch clamp tchnique (whole cell configuration). (A) Membrane current induced by antipeptide antibodies (D16) is not significantly enhanced by IAA. (B) Concentration-dependent anti-peptide (D16) induced membrane current. (C) Effect of antipeptide antibody (D16) following the IAA-induced membrane current. (D) Pre-immune serum (D16) fails to induce a membrane current.*

cytokinin glucosides that are normally mobilizable by the plant host. The ability to hydrolyze N7 and N9 glucosides, which accumulate to high levels in some plant species, increases the substrate pool for this enyzme during pathogenesis, and may simultaneously remove one of the plant's means of compensating for abnormally increased cytokinin levels.

C. HRI:
A region of the maize auxin binding protein (ZmERabp1, ABP) containing many amino acids likely to be present at the active auxin binding site was identified. Antibodies were produced against a synthetic polypeptide embracing this region and recognised all three isoforms of maize ABP on Western blots, as well as ABP homologs in other species, both monocots and dicots, demonstrating that the selected region is highly conserved. IgG fractions from some of these sera (notably D16) were shown at the Gif laboratory (subgroup 1.3) to hyperpolarise protoplast membrane potential in a manner that is precisely characteristic of auxin action. This auxin agonist activity was confirmed at Leiden (subgroup 2.3) and by patch clamp analysis (H. Felle, Gießen, Fig. 1).

The results indicate that the selected peptide does form a functionally significant part of the auxin binding site and strongly suggest that ABP is indeed an auxin receptor. This is the first time that antibodies with agonist activity in plants have been produced. Attempts to produce monoclonal antibodies with similar properties have not been successful, but the reason has been identified and a new screening strategy adopted. Surprisingly, the D16 antibodies were antagonistic to auxin-induced gene expression (Leiden) and the basis of this finding is being investigated. The likely importance of the D16 region of ABP is further emphasized by data now available showing it to be the longest stretch of wholly conserved sequence (six sequences from four species).

Previous work had generated both polyclonal antisera and five monoclonal antibodies (MAC 256-260) to whole ABP. By epitope mappping and fragmentation studies, the epitopes recognised by all these antibodies have been located, providing information for structural and functional analysis (MPI, ISV). All anti-ABP sera, including those from Köln and Bonn (subgroup 2.2) recognise two or three dominant epitopes clustered around the glycosylation site. Two monoclonals (MAC 256, 259) are directed at the endoplasmic reticulum (ER) retention sequence KDEL at the C-terminus of ABP. MAC 257 recognises a region near the N-terminus. Protoplasts of transgenic tobacco plants overexpressing maize ABP (Köln) show an increased sensitivity to auxin in the hyperpolarisation assay (ISV). This increase is blocked by MAC 257, showing that the maize ABP is expressed at the surface of the protoplasts. Most plant ER proteins contain an HDEL, rather than the KDEL retention sequence used by ABP. MAC 256 immunofluorescence and immunogold show a punctuate intracellular distribution in contrast to the generalized HDEL ER staining, suggesting that ABP may be confined to a specific ER compartment. We showed earlier that the ABP C-terminus recognised by MAC 256 undergoes an auxin-induced conformational change, masking the KDEL epitope. This could be of functional significance in allowing ABP to escape from the ER to the plasma membrane, where it mediates membrane potential changes. Such secretion of ABP has been reported by Jones and Herman [Plant Physiol. (1993) 101, 595-606), but their data are incompatible with the known properties of ABP in several respects. Using both wild type and transgenic maize cell suspensions we are trying to define conditions that will make it possible to examine ABP secretion in a meaningful manner.

The ULeeds subgroup has shown that the C-terminal region of ABP may be involved in auxin signal transduction and preliminary data using MAC 256 support this (collaboration with M. Blatt, Wye, in both cases). For future analysis we are preparing antisera against the C-terminal decapeptide. Inclusion of AlF_4^- during maize microsome isolation enhances auxin binding, suggesting a possible role for GTP-binding proteins in auxin signalling. Crosslinking studies were not able to identify any clear interactions of ABP with G proteins or other possible cellular partners. Since most ABP is in the ER, which is a calcium store, ABP may mediate auxin-induced changes in intracellular calcium and processes involved in ER calcium import and export are being studied. Calcium-dependent ATPase of maize shoot ER showing good calmodulin stimulation has been characterized with respect to both hydrolytic and Ca^{2+} transport activities. Some evidence of *in vivo* modulation by auxin has been obtained and liposomal reconstitution methods have been developed for *in vitro* work. We have purified the ER Ca^{2+}-binding protein calreticulin from maize shoots, confirmed authenticity by N-terminal sequencing and raised antisera. These will be used for immunocytology to examine whether the discrete distribution of ABP reflects preferential localization in specialised Ca^{2+} storage ER.

D. ISV:

The auxin-induced modifications of the transmembrane potential difference of protoplasts isolated from tobacco leaves were antagonized by polyclonal antibodies raised to ZmERabp1 (maize abp1), whereas a polyclonal antibody raised to a synthetic peptide reproducing the putative auxin-binding domain of abp1 exhibited auxin antagonist activity on this electrical response. This set of data confirms previous evidence that the recognition of the auxin signal occurs at the outer face of plasma membrane, and suggests that tobacco auxin responsive proteins (tARPs) involved in the auxin response are related to maize abp1. Further investigations of a possible functional role of ZmERabp1 at the plasma membrane were conducted on protoplasts from tobacco plants overexpressing this maize protein (pBAR plants, cooperation with MPI). Such protoplasts exhibited an increased auxin sensitivity of the electrical response by a factor of 30. This increase could be selectively nullified by a monoclonal antibody to maize abp1 (MAC257) which showed no detectable effect on the response of untransformed tobacco protoplasts, i.e. which did not recognize tARPs. This suggests that the increased sensitivity of pBAR protoplasts is due to the expression of maize abp1 at the plasma membrane, and may reflect the routing of a fraction of endogenous abp1 from the ER to the plasma membrane surface.

Tobacco protoplasts transformed by the *rolB* gene from *Agrobacterium rhizogenes* were further analyzed in terms of kinetics of both *rolB* mRNA expression and sensitivity of the membrane response, following auxin application. In conditions where *rolB* mRNA were accumulated and a maximal increase in sensitivity was observed, *rolB*-transformed protoplasts did not differ form normal ones in their ability to accumulate and metabolize exogenous auxins, but showed important changes in their reactivity to the anti-ZmERabp1 antibodies. This supports to the idea that the increased auxin sensitivity of *rolB*-transformed protoplasts could follow from alterations in the reception-transduction of the auxin signal.

To identify the molecular mechanisms of the electrical response of protoplasts to auxin, ion channels at the plasma membrane of protoplasts from tobacco cell suspensions were studied by the patch-clamp technique. In the whole-cell configuration, a voltage-dependent anion current was identified, with a current maximum around -100 mV. The voltage dependence of the whole cell current reflecting

the fast deactivation of the current at potentials <-100 mV was observed only in the presence of internal ATP or when ATP was replaced by the protein phosphatase inhibitor okadaic acid, and was suppressed by staurosporine, a protein kinase inhibitor. This suggests that protein phosphorylation may be involved in regulating the activity of the tobacco suspension anion channel (TSAC). The active auxin 1-NAA caused a time- and concentration-dependent shift of the activation potential of TSAC. In addition, TSAC reacted to the auxin agonist antibody D16 providing evidence for the recognition of the auxin signal at the outer face of the plasma membrane. Further work is needed to assess if the auxin modulation of the anion channel involves a direct interaction between the hormone and the channel protein(s) or the activation of the perception chain including abp1 homologues.

Work is in progress, on the basis of the two-component model we proposed for the organization of auxin perception units at the plasma membrane, to identify both the auxin binding unit and the transmembrane protein in tobacco. cDNA cloning of abp1 homologues in tobacco was achieved via PCR amplification of fragments encompassing conserved regions of abp1. To fish for the transmembrane protein, we used the ability of exogenous maize abp1 to associate in a functional way with the plasma membrane of tobacco protoplasts. Biotinylated abp1 was shown to bind predominantly to one protein from plasma membrane fractions. This protein is presently being characterized.

E.. RUG:
Cell division is an integral part of growth and development. In recent years, it became evident that the basic mechanisms controlling cell division is evolutionary conserved, including in plants. In short, a specific Ser/Thr protein kinase controls the progression of the cell cycle at the G1/S and G2/N boundaries, respectively (Fig. 2). The Ser/Thr protein kinase consists of a 33- to 34-kDa catalytic subunit, now generally referred to as Cdk, and a regulatory subunit, Cyclin. In several organisms the existence of various Cdk/Cyclin complexes has been demonstrated. Until now, the Laboratorium voor Genetica (Gent, Belgium) has cloned and characterized two *cdk* genes, called *cdc2a* and *cdc2b*, and five cyclins, *cyc1, cyc2a, cyc2b, cyc3a,* and *cyc3b*. RNA hybridizations on sorted nuclei revealed that cyc1 is mainly expressed in the G_2 phase whereas *cyc3b* appears to be a G_1 cyclin. For *cyc2a* no pronounced difference in expression levels between G_1 and G_2 nuclei could be detected. Relative amounts of *cdc2a/cdc2b* steady-state transcripts in different plant organs were determined by RT-PCR. *cdc2a* steady-state mRNA levels are the highest in all organs of *Arabidopsis thaliana* except for cell suspensions and flowers where comparable amounts of *cdc2a* and *cdc2b* transcripts could be found. cdc2b mRNA was almost undetectable in stem and very low in siliques and roots.

In both budding yeast (*Saccharomyces cerevisiae*) and fission yeast (*Schizosaccharomyces pombe*) several dominant negative mutants of the $p34^{cdc2}$ (called CDC28p in *S. cerevisiae*) have been identified. In general, two classes of mutant kinases were found. The first type leads to a more active kinase activity, resulting into precocious cell division and a smaller average cell size ('Wee' phenotype). One such mutation is $Tyr_{15} \rightarrow Phe_{15}$. This tyrosine at position 15 becomes phosphorylated, thereby keeping the kinase in an inactive form. Only when the *cdc25*-phosphatase dephosphorylates the Tyr_{15} residue, the kinase becomes active and starts to phosphorylate numerous substrates.

A second class of mutant $p34^{cdc2}$ proteins, when present in a wild-type yeast strain, will completely arrest cell division, without arresting cell growth.

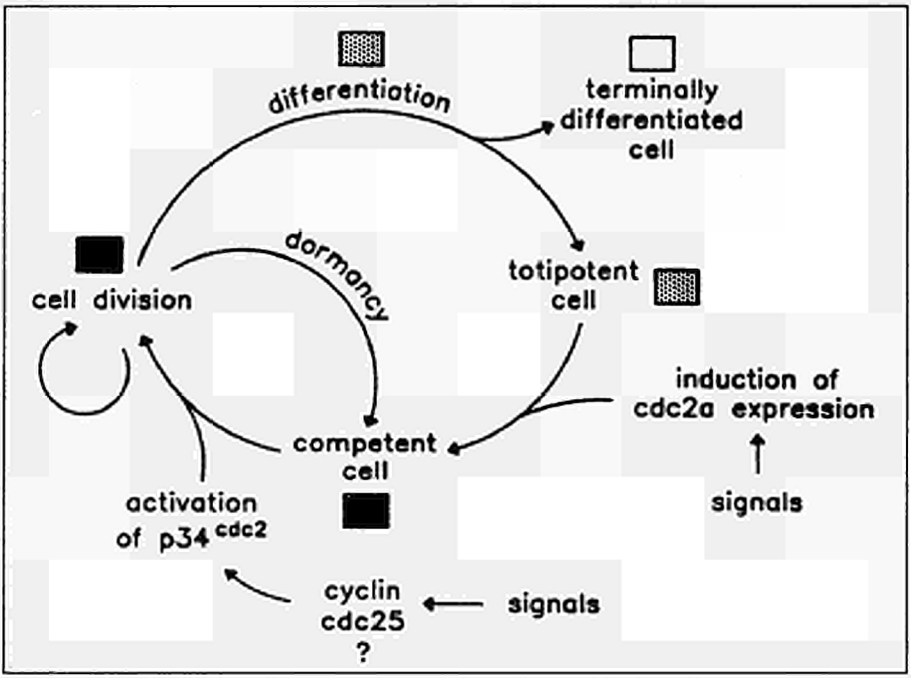


Fig. 2. Schematic representation of a highly simplified model for the regulation of cdc2a *expression during plant development. Black and stippled squares represent cells with high levels of* cdc2a*; white squares represent cells with no detectable* cdc2a *mRNA.*

We have engineered several dominant negative mutations in the Arabidopsis cdc2a gene. The first mutation is $T_{14}Y_{15} \rightarrow V_{14}F_{15}$ and is thought to lead to a faster cell division. The second mutation changes $D_{147}N_{147}$. The Asp residue at position 147 is conserved in all kinases, and is recently shown to have a role in the activation of the kinase. Subsequently, the coding sequences of the wild-type *cdc2* (WT), the D-N, and the TY-VF mutant forms were cloned under the control of the 35 CaMV promoter, all in a binary vector. By transforming root segments of *Arabidopsis*, transgenic plants expressing the WT and TY-VF *cdc2* genes were obtained, although with a seemingly reduced frequency for the TY-VF mutation. With both constructs, both with a low frequency, abnormal phenotypes such as dwarfing, bushy rosette, reduced apical dominance, abnormal, and asymmetric rosette leaves were observed. When expressed in tobacco no phenotypes were observed. In contrast, we were unable to recover any transgenic *Arabidopsis* plant expressing the mutant D-N *cdc2a* gene, suggesting that the presence of the D-N mutant proteins completely abolish cell division. In tobacco, a few transgenic plants were isolated which express the D-N mutant protein. Interestingly, the plants developed normally, although they remained smaller. The number of leaves are identical, compared to untransformed plants, but they have a reduced size. The flowers are considerable smaller. In the various flower organs, a significantly smaller cell number was found. However, the cell size in these organs is as in control plants. In root, stem and leaves of the D-N transgenic plants, cell size are comparatively very large, and as a result the number of cells building up the organs is also considerably smaller. Nevertheless, these plants develop normally. These results

clearly demonstrate that cell division and morphogenesis are two independent events. It furthermore illustrates the extraordinary plasticity plants have. Even with a large reduction in cell number and a large increase in cell size, the intrinsic plant developmental program determines that all organs are generated.

The various *cdc2a* mutant genes were also expressed in *Arabidopsis* under the control of the seed storage protein 2S-1 gene promoter, directing expression only during late emryogenesis. As expected, normal transformation frequencies were observed for all constructs. Transgenic plants expressing too high levels of the D-N *cdc2a* mutant genes segregate seedlings with abnormal phenotypes such as three cotyledons, additional formation of first leaves, one cotyledon, deformed cotyledons. These results demonstrated that overexpression of the D-N *cdc2* gene interferes, at an undefinite time point, with cell division, thereby affecting further development.

F. ULeeds:
At least one route of auxin signalling is known to be mediated by the auxin binding protein (ERabp, ABP) which is surface localised and thought to couple to a transmembrane docking protein. Since G-protein linked signalling represents one of the major routes by which external signals are transduced, and there is strong evidence for the presence of G-proteins within plant cells. We sought to isolate and characterise these proteins and to identify likely receptors and effectors.

1. ***Plant G-proteins, immunological evidence:***
G-protein α-subunits display a high degree of conservation at the primary structural level, with domains being conserved across virtually all G-proteins, e.g. G_{α}common, or specific to specific subfamilies, e.g. G_{α}i and G_{α}Ara (*Arabidopsis* G_{α}). Fig. 3 shows a typical blot of *Pisum* microsomal membrane proteins probed with antisera against synthetic peptides whose sequences correspond to the above domains. Of particular importance is the demonstration of specific immune recognition by preabsorption of the antisera with the peptides to which they were raised (tracks 2,4,6 in Fig. 3).

2. ***Plant G-proteins, guanine nucleotide binding studies:***
An important attribute of G-proteins is that the G_{α} subunits bind GTP and its analogues specifically and with high affinity (Kd $\leftarrow 10^{-8}$). Table I shows that plasma membranes from a number of species display apparent binding constants of this order for the non-hydolysable GTPγS; however, the same membranes also showed binding sites for this nucleotide of somewhat lower affinity. Given the confounding effects of GTPγS-binding to the enzyme nucleoside diphosphate kinase (NDK — see below) the binding of GTPγS cannot necessarily be used as a reliable index of G-protein presence or isolation.

Table I: GTP S-binding to plasma membranes in the presence of 0,01% (v/v) Lubrol PX

	Site 1		*Site 2*	
Membranes	$Kd_{app.}$ *(nM)*	B_{max} *(pMol/mg)*	$K_{dapp.}$ *(nM)*	B_{max} *(pMol/mg)*
Zea	0.7	7	43	66
Pisum	2.6	8	308	174
Nicotiana	8.9	44	298	118

protein. In addition, the interaction of ABP with the proposed docking (DP) must take place between surface localised domains of both proteins. Accordingly, oligopeptides corresponding to predicted solvent exposed domains (Table II) of the major maize ABP (ZmERabp1) were synthesized and used in affinity isolation, guanine nucleotide binding, electrophysiological and ion imaging studies in order to probe events downstream of the ABP within the auxin signalling cascade.

Affinity isolation: A peptide corresponding to the C-terminal 16 residues of ZmERabp1, i.e. P_{148}-L_{163} (peptide A6.2) was immobilised via its N-terminus to prepare an affinity resin for the isolation of the DP. Subsequently, microsomal membranes prepared from etiolated *Zea mays* seedlings essentially were solubilised using a combination of 30 mM MEGA9 and 1% (w/v) cholate and subsequently applied, with recirculation to the A6.2 peptide affinity column. Subsequent elution with 0 to 1M NaCl resulted in the specific elution of an approximately 35 kDa polypeptide. Currently we are working to isolate sufficient of the putative 35 kDa DP to permit N-terminal microsequencing.

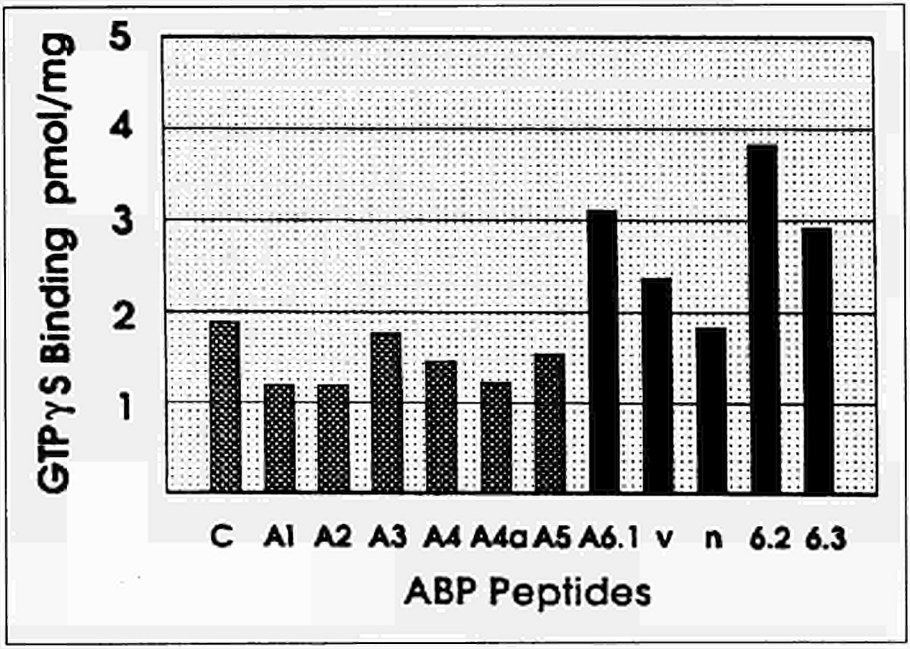


Fig. 5. Effect of ZmERabp1 peptides on GTPγS binding Zea *microsomal membranes.*

Guanine nucleotide binding: The peptides A1 to A6.1 plus the A6 variant peptides listed in Table III were tested for their ability to modulate the binding of [^{35}S]GTPγS to *Zea mays* microsomal membranes. Binding was monitored via rapid filtration through nitrocellulose discs and the non-specific component of binding determined by inclusion of excess GTP within the reaction mix. The effect of 1 μM of each peptide on binding is shown in Fig. 6; only those peptides based on the C-terminal sequence of ZmERabp1 were active in stimulating GTPγS binding. Interestingly, peptide A6.3 which lacked the C-terminal KDEL sequence was effective whilst peptide A6.1v, in which KDEL was present but the rest of the peptide of inverted sequence, was ineffective.

protein. In addition, the interaction of ABP with the proposed docking (DP) must take place between surface localised domains of both proteins. Accordingly, oligopeptides corresponding to predicted solvent exposed domains (Table II) of the major maize ABP (ZmERabp1) were synthesized and used in affinity isolation, guanine nucleotide binding, electrophysiological and ion imaging studies in order to probe events downstream of the ABP within the auxin signalling cascade.

Affinity isolation: A peptide corresponding to the C-terminal 16 residues of ZmERabp1, i.e. P_{148}-L_{163} (peptide A6.2) was immobilised via its N-terminus to prepare an affinity resin for the isolation of the DP. Subsequently, microsomal membranes prepared from etiolated *Zea mays* seedlings essentially were solubilised using a combination of 30 mM MEGA9 and 1% (w/v) cholate and subsequently applied, with recirculation to the A6.2 peptide affinity column. Subsequent elution with 0 to 1M NaCl resulted in the specific elution of an approximately 35 kDa polypeptide. Currently we are working to isolate sufficient of the putative 35 kDa DP to permit N-terminal microsequencing.

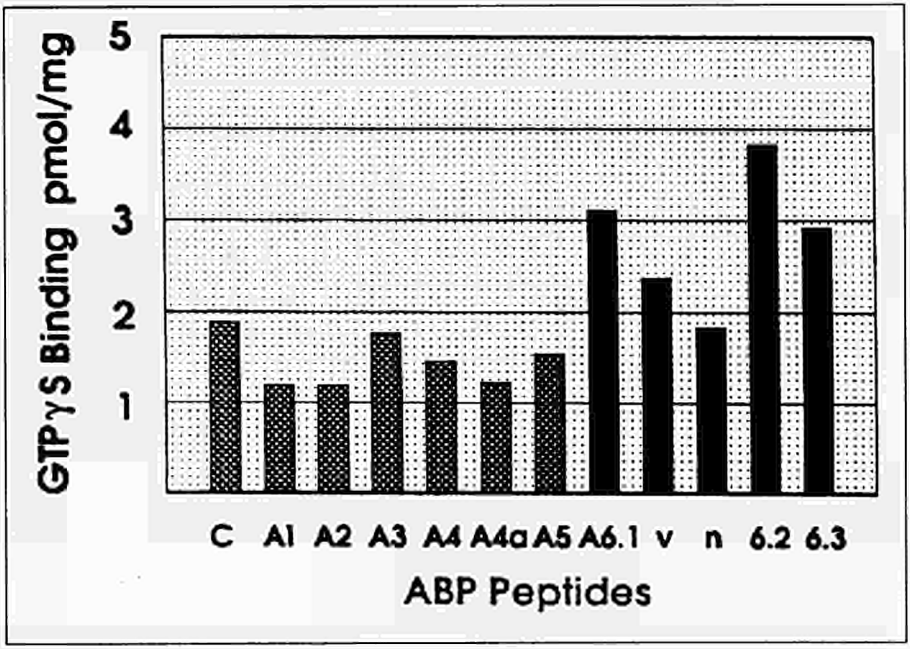


Fig. 5. Effect of ZmERabp1 peptides on GTPγS binding Zea *microsomal membranes.*

Guanine nucleotide binding: The peptides A1 to A6.1 plus the A6 variant peptides listed in Table III were tested for their ability to modulate the binding of [^{35}S]GTPγS to *Zea mays* microsomal membranes. Binding was monitored via rapid filtration through nitrocellulose discs and the non-specific component of binding determined by inclusion of excess GTP within the reaction mix. The effect of 1 μM of each peptide on binding is shown in Fig. 6; only those peptides based on the C-terminal sequence of ZmERabp1 were active in stimulating GTPγS binding. Interestingly, peptide A6.3 which lacked the C-terminal KDEL sequence was effective whilst peptide A6.1v, in which KDEL was present but the rest of the peptide of inverted sequence, was ineffective.

Table II: Peptides corresponding to the predicted external loops of *ZmERadp1*. Numbers indicate the position of the initial and final residues within the ZmERabp1 sequence

Peptide	*Sequence*	*Peptide*	*Sequence*
A1	$_{10}$RDLSQMPQSSYG$_{21}$	A6.1	$_{152}$DEDCFEAAKDEL$_{163}$
A2	$_{32}$GALNHGMKEVE$_{42}$	A6.2	$_{148}$PFVWDEDCFEAAKDEL$_{163}$
A3	$_{51}$GQRTPIHRHSCEE$_{63}$	A6.3	$_{148}$PFVWDEDCFEAA$_{159}$
A4	$_{80}$SLKYPGQPQEIPF$_{92}$	A6.1v	$_{155}$CFEAAKDEL$_{163}$
A4a	$_{92}$FFQNTTFSIPVSD$_{104}$	A6.1n	$_{159}$AAEFCDEDKDEL$_{163}$
A5	$_{104}$DPHQVWNSDEHEDL$_{117}$	152	

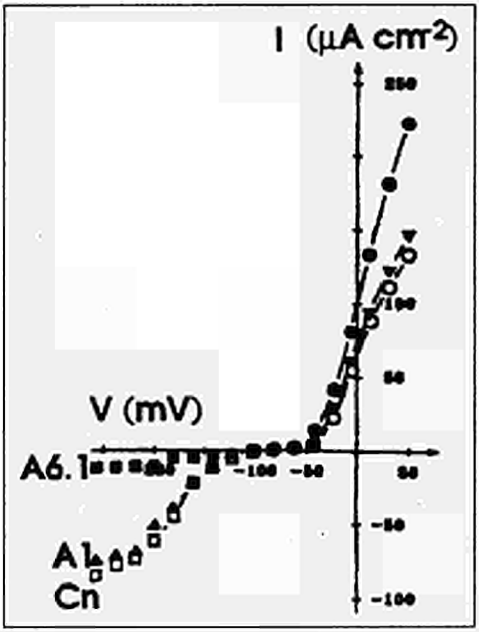


Fig. 6: Effect of ZmERadp1 peptides on K^+ channels in Vicia *guard cells.*

Electrophysiological studies: the effects of externally applied peptides A1 to A6.1 on inwardly rectified Ca^{2+} and H^+ dependent K^+ channels in intact *Vicia* guard cells was tested. Again, peptides A1 to A5 were without effect whereas A6.1 acted as a potent inhibitor (Fig. 6) with I_{50} of about 5 µM. In parallel ion imaging studies, A6.1 elevated intracellular Ca^{2+} and H^+ whereas the other peptides were again ineffective.

MAJOR SCIENTIFIC BREAKTHROUGHS

Several novel proteins have been tagged by photoaffinity labeling. Tagging allowed purification, partial sequencing of peptides and subsequent gene isolation. Current evidence indicates that enzymes that can hydrolyze phytohormone conjugates play a role in initiating important regulatory pathways. They provide invaluable tools for studying the mechanisms underlying growth and development in plants. Auxin

agonist antibodies directed against the highly conserved auxin-binding domain can be used as probes for other auxin-binding proteins (e.g. nuclear cytoplasmic receptors) which can be expected to share this domain. These antibodies, together with antibodies MAC 256/259 and new polyclonal sera against the other known functional domain (C-terminus) will be valuable in exploring auxin signal transduction by electrophysiological methods. Membrane potential changes have been identified which are in line with a modulation of anion channels. The identification of the plasma membrane receptor(s) involved is already in progress. The functional characterization of the transduction steps will then be envisaged taking advantage of the new possibilities offered by the patch-clamp technique. Cell division control genes have been cloned and analyzed. Several dominant mutants were engineered and analysis of transformed plants showed an uncoupling of cell division. Even with a large reduction in cell number the intrinsic developmental program determined all organs. First evidence for G protein coupled receptors was obtained and a detailed biochemical analysis of G proteins was achieved. Most importantly evidence was obtained using bioactive synthetic peptides that ZmERabp1 is an auxin receptor that affects potassium transport processes in stomata.

MAJOR COOPERATIVE LINKS

A series of visits between the different labs was organized for informal discussion, project management and collaborative work. Many tools have been build that were distributed within the group as well as to other groups of the T-program. Anti-ABP antibodies were provided to MPI, HRI, RUG and to other T-project labs. Transgenic tobacco were distributed from MPI (HRI, ISV), anti-Gα antibodies from ULeeds (MPI, HRI, ISV).

PUBLICATIONS

Joint publications

BARBIER-BRYGOO, H., EPHRITHIKINE, G., KLÄMBT, D., MAUREL, C., PALME, K., SCHELL, J., GUERN, J. (1991) Perception of the auxin signal at the plasmalemma of tobacco protoplasts. *Plant Journal* **1**, 83-93.

BERGOUNIOUX, C., PERGENNES, C., HEMERLY, A., QIN, L.X., SARDA, C., INZE, D., and GADAL, P. (1992) A *cdc2* gene of *Petunia hybrida* is differentially expressed in leaves, protoplasts and during various cell cycle phases. *Plant Mol. Biol.* **20**, 1121-1130.

HEMERLY, A., BERGOUNIOUX, C., VAN MONTAGU, M., INZÉ, D., and FERREIRA, P. (1992). Genes regulating the plant cell cycle: isolation of a mitotic-like cyclin from *Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA* **89**, 3295-3299.

RÜCK, A., PALME, K., VENIS, M.A., NAPIER, R.M., and FELLE, H.(1993) Patch- clamp anlysis establishes a role for an auxin binding protein in the auxin stimulation of plasma membrane current in *Zea mays* protoplasts. *Plant J.* **4**, 41-46.

VENIS, M.A., NAPIER, R.M., BARBIER-BRYGOO, H., MAUREL, C., PERROT-RECHENMANN, C. and GUERN, J. Antibodies to a peptide from the maize auxin-binding protein have auxin agonist activity. *Proc. Natl. Acad. Sci. USA* **89**, 7208-7212.

Individual publications

ANUNTALABHOCHAI, S., TERRYN, N., VAN MONTAGU, M., and INZÉ, D. (1991) Molecular characterization of an *Arabidopsis thaliana* cDNA encoding a small GTP-binding protein, Rha1. *Plant J.* **1**, 167-174.

BRZOBOHATY, B., MOORE, I., KRISTOFFERSEN, P.,BAKO, L., CAMPOS, N., SCHELL, J., and PALME, K. (1993) Release of active cytokinin by a β-glucosidase localized to the maize root meristem. *Science* **262**, 1051-1054.

CAMPOS, N., FELDWISCH, J., ZETTL, R., BOLAND, W., SCHELL, J., AND PALME, K. (1991) Identification of auxin-binding proteins using an improved assay for photoaffinity labeling with 5-N_3-[7-H]-indole-3-acetic acid. *Technique* **3**, 69-75.

CAMPOS, N., BAKO, L., FELDWISCH, J., SCHELL, J., and PALME, K. (1992) A protein from maize labeled with azido-IAA has novel β-glucosidase activity. *Plant J.* **2**, 675- 684.
CLARKSON, J., WHITE, I.R. AND MILLNER, P.A. (1991) Specific immune detection and partial purification of G-proteins from *A. thaliana. Biochem. Soc. Trans.* **20**, 9S.
EPHRTIKHINE, G., PRADIER, J.-M., and GUERN, J. (1993) Complexity of GTPγS binding to tobacco plasma membranes. *Plant Physiol. Biochem.*, **31** , 573-584.
FELDWISCH, J., ZETTL, R., HESSE, F., SCHELL, J., AND PALME, K. (1992) Novel auxin binding proteins are localised to the plasma membrane of maize coleoptile cells: Identification by photoaffinity labeling and purification of a 23 kDa polypeptide. *Proc. Natl. Acad. Sci. USA* **89**, 475-479.
FELLE, H., PETERS, W., and PALME, K. (1991) The electrical response of maize to auxins. *Biochim. Biophys. Acta* **1064**, 199-204.
FERREIRA PCG, HEMERLY AAS, VAN MONTAGU M, and INZE D (1993) A protein phosphatase 1 from *Arabidopsis thaliana* restores temperature sensitivity of a *Schizosaccharomyces pombe cdc25^{ts}/wee1^{-}* double mutant. *Plant J* **4**, 81-87.
FERREIRA, P.C.G., HEMERLY, A.S., VILLARROEL, R., VAN MONTAGU, M., and INZÉ, D. (1991). The *Arabidopsis* functional homolog of the p34^{cdc2} protein kinase. *Plant Cell* **3**, 531-540.
FINAN, P.M., WHITE, I.R., FINDLAY, J.B.C. AND MILLNER, P.A. (1991). Identification of nucleoside diphosphate kinase from pea microsomal membranes. *Biochem. Soc. Trans.* **20**, 10S.
FINAN, P., WHITE, I.R., REDPATH, S.H., FINDLAY, B.C., and MILLNER, P.A. (1994) Molecular cloning, sequence determination and heterologous expression of nucleoside diphosphate kinase from *Pisum sativum. Plant Mol. Biol.* **25**, 59-67.
GROARKE, A., WEBSTER, J., WHITE, I.R., and MILLNER, P.A. (1993) Identification of a membrane receptor for the auxin binding protein. *Biochem. Soc. Trans.* in press.
HEMERLY AAS, FERREIRA PCG, ENGLER J, ENGLER G, INZE D, and VAN MONTAGU M (1993) The control of cell cycle in *Arabidopsis* plant cell cultures. *J. Plant Res.* **3**, 51-56.
HEMERLY AAS, FERREIRA PCG, de ALMEIDA ENGLER J, VAN MONTAGU M, ENGLER G, and INZE D (1993) *cdc2*a expression in *Arabidopsis* is linked with competence for cell division. *Plant Cell* **5**, 1711-1723.
HESSE, T., and PALME, K. (1994) The search for phytohormone receptors: A role for auxin binding proteins? In: *Biochemical Mechanisms Involved in Growth Regulation*. Chiatante, D., Gallon, J., Smith, C., Zocchi, (eds.). Oxford University Press, pp. 103-110.
INZÉ, D., FERREIRA, P., HEMERLY, A., and VAN MONTAGU, M. (1992). Control of cell division in plants. *Biochem. Soc. Trans.* **20**, 80-84.
KÖRBER, H., STRIZHOV, N., STAIGER, D., FELDWISCH, J., OLSSON, O., SANDBERG, G., PALME, K., SCHELL, J., KONCZ, C. (1991) T-DNA gene 5 of *Agrobacterium tumefaciens* modulates auxin response by autoregulated synthesis of a growth hormone antagonist in plants. *EMBO J.* **10**, 3983-3991.
MAUREL, C., BARBIER-BRYGOO, H., SPENA, A., TEMPE, J., and GUERN, J. (1991) Single *rol* genes from *Agrobacterium rhizogenes* TL-DNA alter some of the cellular responses to auxin in *Nicotiana tabacum. Plant Physiol.* **97**, 212-216.
MOORE, I., FELDWISCH, J., CAMPOS, N., ZETTL, R., BRZOBOHATY, B., BAKO, L., SCHELL, J., PALME, K. (1991) Auxin-binding proteins of *Zea mays* identified by photoaffinity labelling. *Biochem. Soc. Transact.* **20**, 70-73.
NAPIER, R and VENIS MA (1991) From auxin binding protein to plant hormon receptor. *TIBS* **16**, 72-75.
NAPIER, R.M., and VENIS, M.A. (1992) Epitope mapping reveals conserved regions of an auxin binding protein. *Biochem. J.* **284**, 8441-8445.
NAPIER, R.M., and VENIS, M.A. (1992) The auxin receptor: structure and distribution. In *Long Ashton Symposium 'Transport and receptor proteins of plant membranes'* (Eds. D.T. Clarkson and D. Cooke). Vol. 12. Plenum Press, New York, pp. 169-177.
NAPIER, R.M., FOWKE, L.C., HAWES, C.R., LEWIS, M., and PELHAM, H.R.B. (1992) Immunological evidence that plants use both HDEL and KDEL for tageting proteins to the endoplasmic reticulum. *J. Cell Sci.* **102**, 261-271.

NITSCHKE, K., FLEIG, U., SCHELL, J., and PALME, K. (1992) Complementation of the cs dis 2-11 cell cycle mutant of *Schizosaccharomyces pombe* by a protein phosphatase from *Arabidopsis thaliana. EMBO J.* **11**, 1327-1333.
PALME, K., HESSE, T., MOORE, I., CAMPOS, N., FELDWISCH, J., GARBERS, C., HESSE, F., and J. SCHELL (1991) Hormonal regulation of plant growth: The role of auxin perception. *Mech. Devel.* **33**, 97-106.
PALME, K., and SCHELL, J. (1991) Auxin receptors take shape. *Current Biology* **4**, 137-139.
PALME, K. (1992) The relevance of eucaryotic signalling models to understand the perception of auxin. In *Control of Plant Gene Expression*. Verma, D.P.S. (ed.) Telford Press, in press.
PALME, K., DIEFENTHAL, T., VINGRON, M., SANDER, C., and SCHELL, J. (1992) Molecular cloning and structural analysis of genes from *Zea mays* (L.) coding for members of the *ras*-related *ypt* gene family. *Proc. Natl. Acad. Sci. USA* **89**, 787-791.
PALME, K., FELDWISCH, J., PETERS, W.F., SCHELL, J., ZETTL, R., CAMPOS, N., and FELLE, H. (1992) Auxin-binding proteins are located in the ER and in the plasma membrane: Identification by photoaffinity-labelling and characterization. In *Progress in Plant Growth Regulation.* C.M. Karssen, L.C. Van Loon, D. Vreugdenhill (eds.). Kluwer Academic Publishers, pp. 73-81.
PALME, K., HESSE, T., CAMPOS, N., GARBERS, C., YANOFSKY, M.F., SCHELL, J. (1992) The auxin-binding protein from *Arabidopsis thaliana* is located on chromosome 4. *Plant Cell* **4**, 193-201.
PALME, K. (1992) The relevance of eucaryotic signalling models to understand the perception of auxin. In *Control of Plant Gene Expression.* Verma, D.P.S. (ed.) Telford Press, pp. 33-50.
RICART, C.A., WEBSTER, J., WHITE, I.R., and MILLNER, P. (1993) Specific immunodetection of and purification of G-proteins from *Sorghum bicolor. Biochem. Soc. Trans.* in press.
SCHELL, J., KONCZ, C., SPENA, A., and PALME, K. (1992) Characterization of genes affecting plant growth and development. In Harnessing Biotechnology for the 21st Century. *Ninth International Biotechnology Symposium.* Ladish, M.R., Bose, A. (eds.). American Chemical Society Press, Washington. **116**, pp. 496-499.
TERRYN, N., ANUNTALABHOCHAI, S., VAN MONTAGU, M., and INZE, D. (1992) Analysis of a Nicotiana plumbaginifolia cDNA encoding a novel small GTP-binding protein. *FEBS Lett.* **299**, 287-290.
TERRYN, N., VAN MONTAGU, and INZE, D. (1993) GTP-binding proteins in plants. *Plant Mol. Biol.* **22**, 143-152.
TERRYN N, ARIAS, M.B., ENGLER, G., TIRE, C., VILLAROEL, R., VAN MONTAGU, M., INZE, D. (1993) *rha*, a gene encoding a small GTP binding protein from *Arabidopsis*, is expressed primarily in developing guard cells. *Plant Cell* **5**, 1761-1769.
THIEL, G., BLATT, M.R., FRICKER, M.D., WHITE, I.R., MILLNER, P.A. (1993) Modulation of K^+ channels in *Vicia* stomatal guard cells by peptide homologs to the auxin-binding protein C-terminus. *Proc. Natl. Acad. Sci.* **90**, 11490-11497.
THOMAS, E.W., CASEY, M.M., and VENIS, M.A. (1993) Synthesis and auxin activity of substituted 5-phenylazo- and 5-styryl-1-naphthaleneacetic acids. *Phytochemistry* **32**, 533- 534.
VENIS, M.A., THOMAS, M.W., BARBIER-BRYGOO, H., EPHRITIKHINE, G., and GUERN, J. (1990) Impermeant auxin analogues have auxin activities. *Planta* **182**, 232-235.
VENIS, M.A. and NAPIER, R.M. (1992). Plant hormone receptors: past, present and future. *Biochem. Soc. Trans.* **20**, 55-59.
VENIS, MA and NAPIER, RM (1994) Auxin-binding proteins. *Critical Rev. Plant Sci.*, (in press).
WHITE, T.R., ZAMRI, I., WISE, A., and MILLNER, P.A. (1993) Use of synthetic peptides to study G-proteins and protein kinases within plant cells. In *SEB Seminar Series 'Post translational modifications in plants'* (Eds. H. Battey, D. Dickinson, and A.M. Hetherington). Vol 53. Cambridge University Press, Cambridge. pp. 91-108.
WHITE, I.R., WISE, A., FINAN, P.M., CLARKSON, J. AND MILLNER, P.A. (1991) GTP-binding proteins in higher plant cells. *Proceedings of 12th Long Ashton Symposium.* (in press).
WHITE, I.R. (1992) In: Transport and Receptor proteins of plant membranes. (Eds. Cooke, DT and Clarkson, DT) Plenum Press, New York, pp 185-192.

WHITE, I.R., WISE, A., and MILLNER, P.A. (1993) Evidence for G-protein-linked receptors in higher plants: stimulation of GTP-gamma-S binding to membrane fractions by the mastoparan analgoue mas7. *Planta* **191**, 285-288.
WHITE I.R., FINAN, P.M., and MILLNER, P.A. (1993) Nucleoside diphosphate kinase associated with *Pisum sativum* microsomal membranes: apparent binding of GTP S at nM concentratons. *J. Pl. Physiol.* **142**, 191-196.
WISE, A. AND MILLNER, P.A. (1991) Evidence for the presence of GTP-binding proteins in tobacco leaf and maize hypocotyl plamalemma. *Biochem. Soc. Trans.* **20**, 7S.
WISE, A., WHITE, I.R., and MILLNER, P.A. (1993) Stimulation of guanosine-5'-O-(thiotriphosphate) binding to higher plant plasma membranes by the mastoparan analogue Mas7. *Biochem. Soc. Trans.* (in press).
ZETTL, R., CAMPOS, N., FELDWISCH, J., BOLAND, W., SCHELL, J. and PALME, K. (1991) 5'-azido-[3,6-$^3H_2$]-naphthylphtalamic acid, a photoactivatable probe for auxin efflux carrier proteins. *Technique* **3**, 151-158.
ZETTL, R., FELDWISCH, J., BOLAND, W., SCHELL, J., and PALME, K. (1992) Azido-[3,6-$^3H_2$]-N-naphthylphtalamic acid, a novel photo-activatable probe for auxin carrier proteins from higher plants: Identification of a 23 kDa protein from maize coleoptile plasma membranes. *Proc. Natl. Acad. Sci. USA* **89**, 480-484.

PARTICIPANT INDEX

KEYWORD INDEX

European Commission

EUR 15777 — Biotechnology Research for Innovation, Development and Growth in Europe
Volume II: Final report

A. Vassarotti, L. Kalbe-Bournonville, J. Baselga de Elorz and Ph. de Taxis du Poët

Luxembourg: Office for Official Publications of the European Communities

1995 — VIII + 766 pp. — 14.8 X 21 cm

Volume I: ISBN 92-826-5407-9
Volume II: ISBN 92-826-5408-7
Volumes I and II: ISBN 92-826-5406-0

Price (excluding VAT) in Luxembourg: Vol. II: ECU 63
Vol. I and II: ECU 69.50

Summary

The European Commission is implementing several priority actions specifically designed to improve the competitiveness of European biotechnology. One of these actions aims at the establishment of a Community network for training and research and has been executed from 1982 to 1989 in the framework of two successive Community programmes: the Biomolecular Engineering Programme (BEP; 15 million ECU; April 1982-March 1986) and the Biotechnology Action Programme (BAP; 75 million ECU; 1985-1989). These programmes aimed therefore at establishing a supporting infrastructure for biotechnology research in Europe and at eliminating of bottlenecks which prevent the exploitation of materials, data and methods from modern biology. The BRIDGE programme (Biotechnology, Research for Innovation, Development and Growth in Europe) covered the period 1990-1994, with a budget of 100 million ECU. The research activities in the BRIDGE programme were conducted through two different types of projects: N-projects and T-projects. Following three successive calls for proposals in 1989 and 1990, 69 N-projects and 7 T-projects were launched during 1990. These encompassed 579 participating organizations (388 for the N-projects and 191 for the T-projects) from 11 Member States and 5 EFTA countries. A 'BRIDGE Catalogue of Contracts' (EUR 14278 — Editor: B. Nieuwenhuis) listing the participating laboratories together with a description of their objectives was published in 1992. On a regular basis the project coordinators submit a progress report covering the scientific results achieved in the various laboratories, the evidence of an increasing integration of work and additional relevant considerations. These contributions were assembled in the 1992 (EUR 14298) and 1993 (EUR 15111) Progress Reports, while the final report is subdivided in two volumes: Volume I, the Catalogue of BRIDGE achievements and Volume II, the final BRIDGE report (EUR 15777.EN).

Venta y suscripciones • Salg og abonnement • Verkauf und Abonnement • Πωλήσεις και συνδρομές
Sales and subscriptions • Vente et abonnements • Vendita e abbonamenti
Verkoop en abonnementen • Venda e assinaturas

BELGIQUE / BELGIË

Moniteur belge / Belgisch staatsblad
Rue de Louvain 42 / Leuvenseweg 42
1000 Bruxelles / 1000 Brussel
Tel. (02) 512 00 26
Fax (02) 511 01 84

Jean De Lannoy
Avenue du Roi 202 / Koningslaan 202
1060 Bruxelles / 1060 Brussel
Tel. (02) 538 51 69
Telex 63220 UNBOOK B
Fax (02) 538 08 41

Autres distributeurs/
Overige verkooppunten:

Librairie européenne/ Europese boekhandel
Rue de la Loi 244/Wetstraat 244
1040 Bruxelles / 1040 Brussel
Tel. (02) 231 04 35
Fax (02) 735 08 60

DANMARK

J. H. Schultz Information A/S
Herstedvang 10-12
2620 Albertslund
Tlf. 43 63 23 00
Fax (Sales) 43 63 19 69
Fax (Management) 43 63 19 49

DEUTSCHLAND

Bundesanzeiger Verlag
Breite Straße 78-80
Postfach 10 05 34
50445 Köln
Tel. (02 21) 20 29-0
Fax (02 21) 202 92 78

GREECE ΕΛΛΑΔΑ

G.C. Eleftheroudakis SA
International Bookstore
Nikis Street 4
10563 Athens
Tel. (01) 322 63 23
Telex 219410 ELEF
Fax 323 98 21

ESPAÑA

Boletín Oficial del Estado
Trafalgar, 27-29
28071 Madrid
Tel. (91) 538 22 95
Fax (91) 538 23 49

Mundi-Prensa Libros, SA
Castelló, 37
28001 Madrid
Tel. (91) 431 33 99 (Libros)
431 32 22 (Suscripciones)
435 36 37 (Dirección)
Telex 49370-MPLI-E
Fax (91) 575 39 98

Sucursal:

Librería Internacional AEDOS
Consejo de Ciento, 391
08009 Barcelona
Tel. (93) 488 34 92
Fax (93) 487 76 59

Llibreria de la Generalitat de Catalunya
Rambla dels Estudis, 118 (Palau Moja)
08002 Barcelona
Tel. (93) 302 68 35
Tel. (93) 302 64 62
Fax (93) 302 12 99

FRANCE

Journal officiel
Service des publications des Communautés européennes
26, rue Desaix
75727 Paris Cedex 15
Tel. (1) 40 58 77 01/31
Fax (1) 40 58 77 00

IRELAND

Government Supplies Agency
4-5 Harcourt Road
Dublin 2
Tel. (1) 66 13 111
Fax (1) 47 80 645

ITALIA

Licosa SpA
Via Duca di Calabria, 1/1
Casella postale 552
50125 Firenze
Tel. (055) 64 54 15
Fax 64 12 57
Telex 570466 LICOSA I

GRAND-DUCHÉ DE LUXEMBOURG

Messageries du livre
5, rue Raiffeisen
2411 Luxembourg
Tel. 40 10 20
Fax 49 06 61

NEDERLAND

SDU Overheidsinformatie
Externe Fondsen
Postbus 20014
2500 EA 's-Gravenhage
Tel. (070) 37 89 880
Fax (070) 37 89 783

PORTUGAL

Imprensa Nacional
Casa da Moeda, EP
Rua D. Francisco Manuel de Melo, 5
1092 Lisboa Codex
Tel. (01) 387 30 02/385 83 25
Fax (01) 384 01 32

Distribuidora de Livros Bertrand, Ld.ª

Grupo Bertrand, SA
Rua das Terras dos Vales, 4-A
Apartado 37
2700 Amadora Codex
Tel. (01) 49 59 050
Telex 15798 BERDIS
Fax 49 60 255

UNITED KINGDOM

HMSO Books (Agency section)
HMSO Publications Centre
51 Nine Elms Lane
London SW8 5DR
Tel. (071) 873 9090
Fax 873 8463
Telex 29 71 138

ÖSTERREICH

Manz'sche Verlags- und Universitätsbuchhandlung
Kohlmarkt 16
1014 Wien
Tel. (1) 531 610
Telex 112 500 BOX A
Fax (1) 531 61-181

SUOMI/FINLAND

Akateeminen Kirjakauppa
Keskuskatu 1
PO Box 218
00381 Helsinki
Tel. (0) 121 41
Fax (0) 121 44 41

NORGE

Narvesen Info Center
Bertrand Narvesens vei 2
PO Box 6125 Etterstad
0602 Oslo 6
Tel. (22) 57 33 00
Telex 79668 NIC N
Fax (22) 68 19 01

SVERIGE

BTJ AB
Traktorvägen 13
22100 Lund
Tel. (046) 18 00 00
Fax (046) 18 01 25
30 79 47

ÍSLAND

BOKABUD LARUSAR BLÖNDAL
Skólavordustig, 2
101 Reykjavik
Tel. 11 56 50
Fax 12 55 60

SCHWEIZ / SUISSE / SVIZZERA

OSEC
Stampfenbachstraße 85
8035 Zürich
Tel. (01) 365 54 49
Fax (01) 365 54 11

BÅLGARIJA

Europress Klassica BK Ltd
66, bd Vitosha
1463 Sofia
Tel./Fax 2 52 74 75

ČESKÁ REPUBLIKA

NIS ČR
Havelkova 22
130 00 Praha 3
Tel. (2) 24 22 94 33
Fax (2) 24 22 14 84

HRVATSKA

Mediatrade
P. Hatza 1
4100 Zagreb
Tel. (041) 430 392

MAGYARORSZÁG

Euro-Info-Service
Honvéd Európa Ház
Margitsziget
1138 Budapest
Tel./Fax 1 111 60 61
1 111 62 16

POLSKA

Business Foundation
ul. Krucza 38/42
00-512 Warszawa
Tel. (2) 621 99 93, 628-28-82
International Fax&Phone
(0-39) 12-00-77

ROMÂNIA

Euromedia
65, Strada Dionisie Lupu
70184 Bucuresti
Tel./Fax 1-31 29 646

RUSSIA

CCEC
9,60-letiya Oktyabrya Avenue
117312 Moscow
Tel./Fax (095) 135 52 27

SLOVAKIA

Slovak Technical Library
Nm. slobody 19
812 23 Bratislava 1
Tel. (7) 5220 452
Fax : (7) 5295 785

CYPRUS

Cyprus Chamber of Commerce and Industry
Chamber Building
38 Grivas Dhigenis Ave
3 Deligiorgis Street
PO Box 1455
Nicosia
Tel. (2) 449500/462312
Fax (2) 458630

MALTA

Miller distributors Ltd
PO Box 25
Malta International Airport
LQA 05 Malta
Tel. 66 44 88
Fax 67 67 99

TÜRKIYE

Pres AS
Istiklal Caddesi 469
80050 Tunel-Istanbul
Tel. 0(212) 252 81 41 - 251 91 96
Fax 0(212) 251 91 97

ISRAEL

ROY International
PO Box 13056
41 Mishmar Hayarden Street
Tel Aviv 61130
Tel. 3 496 108
Fax 3 648 60 39

EGYPT/ MIDDLE EAST

Middle East Observer
41 Sherif St.
Cairo
Tel/Fax 39 39 732

UNITED STATES OF AMERICA / CANADA

UNIPUB
4611-F Assembly Drive
Lanham, MD 20706-4391
Tel. Toll Free (800) 274 4888
Fax (301) 459 0056

CANADA

Subscriptions only
Uniquement abonnements

Renouf Publishing Co. Ltd
1294 Algoma Road
Ottawa, Ontario K1B 3W8
Tel. (613) 741 43 33
Fax (613) 741 54 39
Telex 0534783

AUSTRALIA

Hunter Publications
58A Gipps Street
Collingwood
Victoria 3066
Tel. (3) 417 5361
Fax (3) 419 7154

JAPAN

Kinokuniya Company Ltd
17-7 Shinjuku 3-Chome
Shinjuku-ku
Tokyo 160-91
Tel. (03) 3439-0121

Journal Department
PO Box 55 Chitose
Tokyo 156
Tel. (03) 3439-0124

SOUTH-EAST ASIA

Legal Library Services Ltd
Orchard
PO Box 05523
Singapore 9123
Tel. 73 04 24 1
Fax 24 32 47 9

SOUTH AFRICA

Safto
5th Floor, Export House
Cnr Maude & West Streets
Sandton 2146
Tel. (011) 883-3737
Fax (011) 883-6569

AUTRES PAYS
OTHER COUNTRIES
ANDERE LÄNDER

Office des publications officielles des Communautés européennes
2, rue Mercier
2985 Luxembourg
Tel. 499 28-1
Télex PUBOF LU 1324 b
Fax 48 85 73/48 68 17

CG-NB-15777-EN-C

NOTICE TO THE READER

All scientific and technical reports published by the European Commission are announced in the monthly periodical '**euro abstracts**'. For subscription (1 year: ECU 63) please write to the address below.

The BRIDGE Programme (*B*iotechnology, *R*esearch for *I*nnovation, *D*evelopment and *G*rowth in *E*urope) covered the period 1990-94, with a budget of ECU 100 million. The research activities in the BRIDGE programme were implemented via two different types of projects: N-projects and T-projects. Following three successive calls for proposals in 1989 and 1990, 69 N-projects and 7 T-projects were launched during 1990 and 1991. These encompassed 579 participating organizations (388 for the N-projects and 191 for the T-projects) from 11 Member States and five EFTA countries. This final BRIDGE progress report covering the period 1990-1994 has been published in two volumes:

Volume I, the catalogue of BRIDGE achievements, which covers in a very concise way the results obtained in the different projects and also provides a general introduction, as well as the addresses of contractors and other additional information;

Volume II, the final BRIDGE report, covers the entirety of the programme giving for each project the results achieved in the different laboratories, the evidence of increasing integration of work and additional relevant information, thus developing the abridged reports bound together in Volume I.

Price (excluding VAT) in Luxembourg:
Vol. II: ECU 63
Vol. I and II: ECU 69.50

ISBN 92-826-5408-7

OFFICE FOR OFFICIAL PUBLICATIONS
OF THE EUROPEAN COMMUNITIES

Luxembourg