Annual Report 2003

Institute of Biotechnology
Centre of Excellence

Contents

INSTITUTE OF BIOTECHNOLOGY
FOREWORD

The Institute started its activities in 1975 as the All-Union Research Institute of Applied Enzymology. Initially, its activities were focused on development of classical technologies to produce and purify enzymes. Subsequently, in response to developing research environment in Lithuania and worldwide, its priorities have concentrated in the field of modern biology linked to biotechnology and mainly to:

- genetic and molecular studies of restriction-modification phenomenon;
- function of genes in yeast;
- research and development of recombinant biomedical proteins;

Since 1992 the name of the Institute has been changed to the Institute of Biotechnology (IBT), which is a non-profit state research institution.

Structure and management of the Institute are determined by the Director of Institute and the Institute Council, comprising 12 scientists, 8 of which are elected by the Institute's research staff and 4 appointed by the Minister of Science and Education of the Lithuanian Republic for a five-year period. The IBT employs about 110 persons, 77 of which belong to the scientific staff including 13 PhD students.

Scientists in the IBT are encouraged actively to participate in training and educational activities. Students from Vilnius University, Gediminas Technical University, Vytautas Magnus University and Kaunas Technological University gain academic knowledge in the fields of molecular biology, genetic engineering, bioinformatics, biotechnology, immunotechnology.

The IBT was selected as a Centre of Excellence in 2002, paying tribute to the quality of research and allowing to host a number of researchers and graduates.

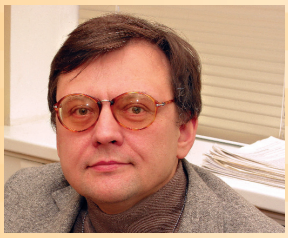
The USA National Academy of Sciences together with Howard Hughes Medical Institute granted the Institute of Biotechnology a right to host *An Intensive Lecture/Laboratory Course on Molecular Interactions of Proteins and Nucleic Acids* with prominent invited speakers and lecturers from Europe and the USA that was held on August 25-September 6, 2003. Principal investigators Habil. Dr. S.Klimašauskas and Dr. V.Šikšnys from the IBT together with Prof. R.Gumport from the University of Illinois at Urbana headed the course. It was attended by 20 students and junior scientists from Bulgaria, Czech Republic, Estonia, Poland, Slovakia, Russia, Hungary and Lithuania. Such courses speed up the pace of research by helping scientists learn about the latest techniques and findings from other laboratories and follow up collaboration.

The IBT had a great influence on the emergence and establishment of several biotechnological spin-offs - UAB Fermentas, UAB Sicor Biotech (formerly named Biotechna) and UAB Biocentras in 1994. Now these companies, in close cooperation with the Institute of Biotechnology, successfully form the developing domestic biotechnological sector in Lithuania.

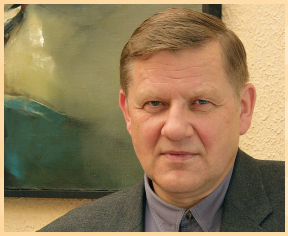
Algimantas Pauliukonis
Director of the Institute

Principal Investigators

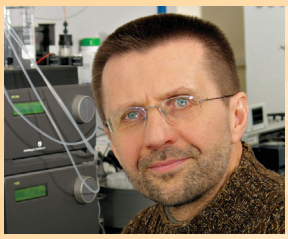
Kęstutis Sasnauskas, PhD, Dr. Habil.
Chief Scientist
phone: 370 5 2602104
fax: 370 5 2602116
e-mail:sasnausk@ibt.lt



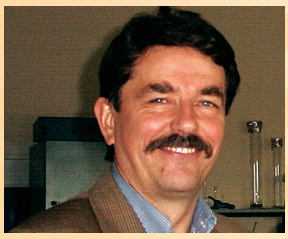
Saulius Klimašauskas, PhD, Dr. Habil.
Chief Scientist
phone: 370 5 2602114
fax: 370 5 2602116
e-mail:klimasau@ibt.lt



Arvydas Janulaitis, PhD, Dr. Habil., Prof.
Chief Scientist
phone: 370 5 2602110
fax: 370 5 2602142
e-mail:janulait@fermentas.com



Virginijus Šikšnys, PhD
Chief Scientist
phone: 3705 2602108
fax: 370 5 2602116
e-mail:siksnys@ibt.lt



Vytautas Naktinis, PhD
Senior Scientist
phone: 3705 2660266
fax: 370 5 2660206
e-mail:Vytautas.Naktinis@sicor.lt

Algimantas Pauliukonis, PhD, Dr. Habil.
Head of the Institute
phone: 370 5 2602103, 370 5 2602106
fax: 370 5 2602116
e-mail:pauliukonis@ibt.lt

Laboratory of Protein-DNA Interactions

Head Dr. Virginijus Šikšnys

Arūnas Lagunavičius, Ph.D.

Saulius. Gražulis, Ph.D.

Česlovas Venclovas, Ph.D.

Elena Manakova, M.Sc.

Giedrius Sasnauskas, M.Sc.

Giedrė Tamulaitienė, M.Sc.

Gintautas Tamulaitis, M.Sc.

Audronė Skirgailienė, M.Sc.

Violeta Stonytė, M.Sc.

Mindaugas Zaremba, M.Sc.

Dmitrijus Golovenko, M.Sc.

Arūnas Šilanskas, M.Sc.

Rasa Sukackaitė, B.Sc.

Ania Tunevič

Restriction enzymes are proteins that are made in bacteria and cut DNA at specific sites. More than 3,000 type II restriction enzymes (there are three classes in all) with nearly 200 differing specificities have been identified to date (Roberts R.J. et. al., 2003). Each recognizes a particular DNA sequence that is four to eight base pairs in length and cleave phosphodiester bonds in the presence of Mg^{2+}, within or close to these sites (Pingoud A. & Jeltsch A., 2001). Restriction endonucleases have now gained widespread application as indispensable tools for the *in vitro* manipulation and cloning of DNA. How these enzymes specifically recognize and cut their DNA targets remains a long-standing question in biology. We focus our studies on the molecular mechanisms by which restriction enzymes achieve their function.

Among the questions being asked are:

- Are all restriction enzymes alike to each other?
- How do the restriction enzymes recognize the particular DNA sequence?
- What are structural and mechanistic relationships within restriction endonuclease family?
- How do the sequence recognition and catalysis are coupled in the function of restriction enzymes?

Answers to these questions are being sought using X-ray crystal structure determination of restriction enzyme-DNA complexes, site-directed mutageneses and biochemical studies to relate structure to function.

Restriction enzyme that acts in the absence of metal ions

Among all of the restriction enzymes found to date, *BfiI* is unique in cleaving DNA without metal ions (Sapranauskas R. et. al., 2000). It recognizes an asymmetric sequence, 5'-ACTGGG-3', and cuts top and bottom strands at fixed positions downstream of this site. We have shown that *BfiI* consists of two physically separate domains, with catalytic and dimerisation functions in the N-terminus and DNA recognition functions in the C-terminus (Zaremba M. et. al., 2004).

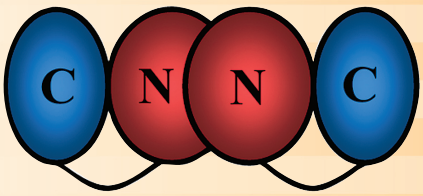


Fig. 1. Modular structure for *BfiI*. In its native state, the *BfiI* endonuclease exists as a dimer. Each subunit is composed of physically separate N-terminal and C-terminal domains. The dimerisation interface involves solely the N-terminal domains and the single active site in the protein is at this interface. The C-terminal domains are responsible for DNA recognition and each can bind one copy of the recognition sequence.

The N-terminal half of *BfiI* is homologous to Nuc, an EDTA-resistant nuclease from *Salmonella typhimurium* that belongs to the phospholipase D superfamily. Like Nuc, the isolated N-terminal domain of *BfiI* is a dimer with non-specific DNA cleavage activity. The dimer has only one active site, at the subunit interface, which engages sequentially first one strand and then the other strand of the DNA (Sapranauskas R. et. al., 2000; Lagunavicius A. et. al., 2003; Sasnauskas G. et. al., 2003). In contrast, the C-terminal domain of BfiI lacks catalytic activity but binds specifically to DNA with the cognate sequence for BfiI (Zaremba M. et.al., 2004). *BfiI* thus carries at its core a dimeric unit capable of cleaving both 5'-3' and 3'-5' strands of the DNA, which is flanked on both sides by DNA-recognition domains. The modular structure of *BfiI* suggests that it may have evolved through the fusion of a DNA-recognition domain to a non-specific Nuc-like nuclease. Elucidation of the crystal structure of *BfiI* is currently in progress.

Principles of restriction enzyme specificity

Due to their unique specificity, Type II restriction enzymes became indispensable tools for the *in vitro* manipulation and cloning of DNA. However, much less is known about how they achieve their specificity. Most of restriction endonucleases share a common structural fold, however, it is still unclear how novel specificities evolve within a conserved structural core. In order to address this question we have focused our efforts on the elucidation of the structural and molecular mechanisms of sequence recognition within

the subgroup of α-class of restriction enzymes. This particular subgroup includes *Cfr*10I, *Bse*634I, *Ngo*MIV, *Kpn*2I, *Ecl*18kI and *Eco*RII restriction endonucleases that recognize related nucleotide sequences N/CCGGN or /CCNGG („/" indicates cleavage position, N stands for the unspecified nucleotide). Moreover, these enzymes share a conserved active site motif that differs from that characteristic for the most of restriction enzymes. Noteworthy, this subgroup includes enzymes of different subtypes (IIP, IIF and IIE) that differ in their quaternary structure. The project thus represents a systematic study focused on the elucidation of the mechanisms of sequence discrimination and structure-function relationships within a group of related restriction enzymes. The experimental approach combines biochemical, mutational and crystallographic studies and should provide us with general principles of specificity of restriction enzymes. Similar studies are performed for other subgroups of restriction endonucleases, e.g., *Sda*I/*Pst*I subfamily of restriction endonucleases (in collaboration with A.Janulaitis).

Over last years we solved crystal structures of *Cfr*10I, *Bse*634I, *Ngo*MIV (Bozic D. et. al., 1996; Deibert M. et. al., 2000, see comments; Grazulis S. et. al., 2002). Recently, we have obtained X-ray quality crystals of *Ecl*18kI and *Kpn*2I restriction endonucleases. Elucidation of their crystal structures will contribute to our understanding of sequence discrimination mechanisms within subfamily of restriction endonucleases.

Structure and function of the tetrameric restriction enzymes

It was thought for a long time that most of Type II restriction endonucleases interact with DNA as homodimers. Our structural studies, however, revealed that restriction enzymes *Cfr*10I (Pu/CCGGPy), *Bse*634I (Pu/CCGGPy) and *Ngo*MIV (G/CCGGC) recognizing closely related nucleotide sequences (shown in brackets, ,/' indicates the cleavage site) are homotetramers (Bozic D. et. al., 1996; Deibert M. et. al., 2000, see comments; Grazulis S. et. al., 2002; Siksnys V. et. al., 1999).

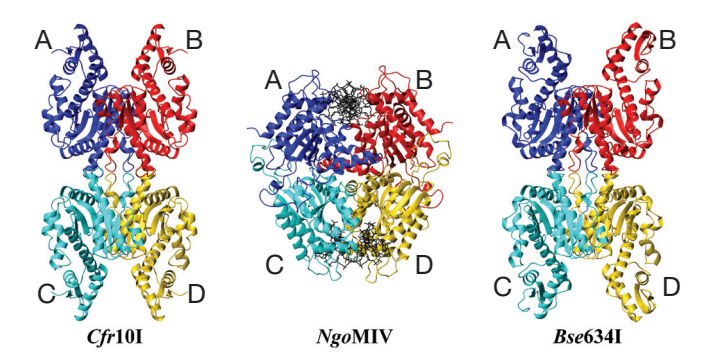


Fig. 2. Tetramers of the *Cfr*10I, *Bse*634I and *Ngo*MIV restriction endonucleases. Monomers are shown as ribbon models in different colours and labelled A, B, C, D, respectively. DNA is shown in a stick representation.

According to the structural data tetramers of restriction enzymes are arranged as a dimer of dimers positioned back-to-back to each other with DNA binding clefts facing opposite directions. Solution experiments demonstrate that tetrameric restriction enzymes *Cfr*10I, *Bse*634I and *Ngo*MIV cleave DNA substrates containing two recognition sites much faster than single-site DNA suggesting that interaction with two sites is of functional importance (Bozic D. et. al., 1996; Deibert M. et. al., 2000, see comments; Grazulis S. et. al. 2002; Siksnys V. et. al., 1999). We have shown, however, that *Bse*634I restriction enzyme can be „cured" of the requirement for the two recognition sites by site-directed mutagenesis. Indeed, we have engineered dimeric variants of *Bse*634I restriction enzyme that cleaved plasmid DNA with a single site nearly with the same efficiency as a two site DNA indicating that interaction with two recognition sites is not absolutely required for the catalytic activity (Zaremba M., Sasnauskas G., Urbanke, C., Grazulis, S. and Siksnys, V., manuscript in preparation). Further experiments are in progress to elucidate correlations between the oligomeric architecture, stability and function of restriction enzymes.

Thermodynamic and biochemical studies of MunI-DNA interactions

Thermodynamic and biochemical characterization of restriction enzyme interactions with cognate and noncognate DNA accompanied by structural studies are necessary for the understanding the mechanisms of sequence recognition. Few years ago we solved the crystal structure of *Mun*I restriction endonuclease in complex with cognate DNA (Deibert M. et. al., 1999). The recognition sequence of *Mun*I (C/AATTG) partially overlaps with that of *Eco*RI (G/AATTC). Crystal structure determination provided us with the structural mechanism of sequence recognition by *Mun*I (Deibert M. et. al., 1999). In order to determine structural-thermodynamic correlations, we employ biochemical and biophysical methods to study *Mun*I restriction endonuclease-DNA interaction in solution. Isothermal titration calorimetry is used to study *Mun*I restriction endonuclease binding to specific and non-specific DNA (Haq I. et. al., 2001). Gel-shift analysis and electron microscopy are employed to characterize protein induced DNA bending in the *Mun*I-DNA complexes. Recently, fluorescent spectroscopy has been applied to study the relationships between *Mun*I oligomeric structure and stability. These studies will contribute to our understanding of structure-function correlation of *Mun*I restriction enzyme and enable to compare mechanisms of sequence discrimination between *Mun*I and *Eco*RI.

Collaboration

Dr. Matthias Bochtler, International Institute of Molecular and Cell Biology, Warsaw
Dr. Dmitry Cherny, MPI Biophysical Chemistry, Gottingen
Prof. Dr. John Ladbury, University College London
Prof. Dr. Steve Halford, University of Bristol, Bristol
Prof. Dr. Robert Huber, MPI of Biochemistry, Martinsried
Prof. Dr. Alfred Pingoud, University of Giessen, Giessen
Dr. Monika Reuter, Institute of Virology, Berlin
Prof. Dr. Claus Urbanke, Hannover Medical School, Hannover

A break with Prof. Dr. Robert Huber F.R.S. (left) from the Max Planck Institute of Biochemistry after his seminar on *Molecular Machines for Protein Degradation* in the main hall of the Institute of Biotechnology in May 2003.

Grants

Howard Hughes Medical Institute
Wellcome Trust
EC Framework 5 Programme
Max Planck's Society
International Quality Network
Lithuanian State Science and Studies Foundation
Lithuanian High Tech Development Programme

Recent publications

Lagunavicius A., Sasnauskas G., Halford S.E., and Siksnys V. (2003). The metal-independent Type IIs restriction enzyme *Bfi*I is a dimer that binds two DNA Sites but has only one catalytic centre. ***J. Mol. Biol.*** 326(4):1051-64.

Sasnauskas G., Halford S.E., and Siksnys V. (2003). How the *Bfi*I restriction enzyme uses one active site to cut two DNA strands. ***Proc. Nat. Acad. Sci. USA*** 100(11):6410-6415

Zaremba M., Urbanke C., Halford S.E. &. Siksnys V. (2004). Generation of the *Bfi*I restriction endonuclease from the fusion of a DNA recognition domain to a non-specific nuclease from the phospholipase D superfamily. ***J. Mol. Biol.*** 336:81-92.

Siksnys V., Grazulis S., Huber R. (2004). Structure and function of the tetrameric restriction enzymes. In: ***Restriction Endonucleases*** (Ed. A. Pingoud), Springer Verlag, p. 237.

Contact: siksnys@ibt.lt

Laboratory of Biological DNA Modification

Head Dr. Habil. Saulius Klimašauskas

Giedrius Vilkaitis, Ph.D.

Saulius Serva, Ph.D.

Eglė Merkienė, M.Sc.

Dalia Daujotytė, M.Sc.

Zdislav Staševskij, M.Sc.

Gražvydas Lukinavičius, M.Sc.

Rūta Gerasimaitė, M.Sc.

Asta Vilčinskienė, M.Sc.

Laura Manelytė, M.Sc.

Giedrė Urbanavičiūtė, B.Sc.

Ala Žilionienė

Enzymatic methylation of nucleobases serves to expand the information content of the genome in organisms ranging from bacteria to mammals. Postreplicative methylation is accomplished by DNA methyltransferases, which catalyze the transfer of a methyl group from cofactor S-adenosyl-l-methionine (AdoMet) yielding 5-methyl cytosine, N4-methylcytosine or N6-methyladenine. DNA methylation is involved in the control of transcription, genomic imprinting, developmental regulation in eukaryotes. Aberrations in cytosine-5 methylation play a role in human genetic disease, and, therefore, the MTases are potent candidate targets for anticancer therapy. MTases are obligatory enzyme constituents of restriction-modification systems found in most bacterial organisms. DNA cytosine-5 MTases from both eukaryotes and prokaryotes comprise a family of homologous enzymes.

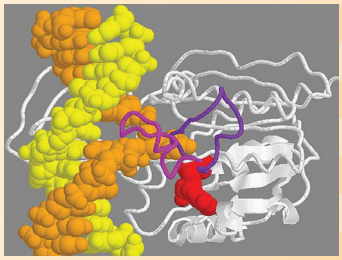


Fig. 1. Action of the *Hha*I methyltransferase. Flipping of the target cytosine out of the DNA (spacefill, yellow and orange) helix is accompanied by a conformational rearrangement of the catalytic loop (violet and magenta). Bound cofactor AdoMet is shown in red.

Besides their diverse biological roles DNA methyltransferases (MTases) are attractive models for studying structural aspects of DNA-protein interaction. Bacterial enzymes recognize an impressive variety (over 200) of short sequences in DNA. As shown first for the *Hha*I MTase, access to the target base that is buried within the stacked double helix is gained in a remarkably elegant manner: by rotating the nucleotide completely out of the DNA helix and into a concave catalytic pocket in the enzyme. This general mechanistic feature named "base-flipping" is shared by numerous other DNA repair and DNA modifying enzymes. Our laboratory studies the mechanistic and structural aspects of DNA methylation using the *Hha*I methyltransferase (M.*Hha*I) from the bacterium *Haemophilus haemolyticus* as the paradigm model system.

Progress during 2002-2003

Kinetic and molecular mechanism of DNA methylation

Enzymatic DNA cytosine-5 methylation is a complex reaction which proceeds via multiple steps such as binding of cofactor AdoMet and substrate DNA, flipping of the target cytosine, conformational rearrangement of the mobile catalytic loop, activation of the target cytosine via formation of a transient covalent bond, methyl transfer and so on. We use mutagenesis, gel-shift binding and chemical foot-printing, steady-state and transient kinetic analysis, fluorescence spectroscopy and *X*-ray diffraction to delineate the elementary steps on the reaction pathway of *Hha*I MTase.

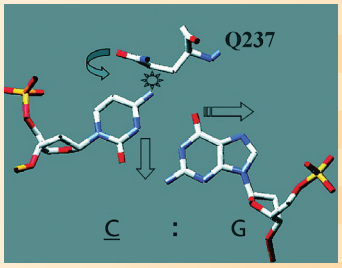


Fig. 2. Molecular interactions between Gln237 and the target base pair in DNA. The steric clash (x) expels the target base into a groove leading to the disruption of base pairing and stacking interactions.

Despite of extensive studies of the *Hha*I cytosine-5 methyltransferase, initial events in the base flipping mechanism of this model enzyme remained elusive. We found that the replacement of the target C:G pair with the 2-aminopurine:T pair in the DNA or certain mutations of Gln237 in the protein severely perturb base flipping, but retain specific DNA binding. Our kinetic analyses and molecular modelling suggest that a steric interaction between the side chain of Gln237 and the target cytosine in B-DNA reduces the energy

barrier for flipping by 3 kcal/mol. Subsequent stabilization of an open state by further 4 kcal/mol is achieved through specific hydrogen bonding of the side chain to the orphan guanine. Deletion mutations in the catalytic loop have no effect on DNA binding and flipping of the 2-aminopurine base. In aggregate, our study for the first time demonstrates that Gln237 plays a key role in actively opening the target C:G pair by a "push-and-bind" mechanism (submitted for publication).

To obtain insights into dynamics of the base flipping we have initiated an NMR structural study of macromolecular interactions and dynamics of the *Hha*I MTase-DNA system in solution. However, prior extensive experience revealed an insufficient solubility of the enzyme (0.15 mM) under reaction conditions. To overcome this problem, three hydrophobic patches on the surface of M.*Hha*I that are not involved in substrate interactions were subjected to site-specific mutagenesis. Residues M51 or V213 were substituted by polar amino acids of a similar size, and/or the C-terminal tetrapeptide FKPY was replaced by a single glycine residue. Two out of six mutants showed improved solubility in initial analyses and were purified to homogeneity using a newly developed procedure. Biochemical studies of the engineered MTases showed that the deletion mutant retained identical DNA binding, base flipping and catalytic properties as the wild type enzyme. In contrast, the engineered enzyme showed a significantly increased solubility (0.35 mM) and long term stability. High-quality [^{15}N,^{1}H] TROSY NMR spectra and ^{15}N spin relaxation times (determined in collaboration with Prof. Thomas Szyperski, SUNY Buffalo, NY, USA) clearly indicate the presence of a monomeric well-folded protein in solution. This work represents the first

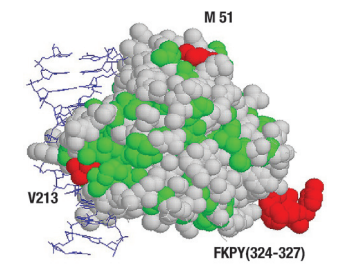


Fig. 3. Spacefill model of the *Hha*I MTase: hydrophobic surface residues are shown in green and the target positions for mutagenesis in red; bound DNA is shown as wireframe.

successfull design of a highly-soluble functionally active variant of a DNA MTase. The modifed *Hha*I MTase can be prepared in high yield as a pure isotope-labeled protein. This now paves the way to detailed structural studies of this model cytosine-5 MTase and its interactions with the ligands employing NMR spectroscopy, stopped-flow fluorescence, isothermal calorimetry and other methods to further our understanding of the mechanism of this physiologically important enzymatic reaction in DNA (ref.)

Correlating structure and spectroscopy: 2-aminopurine fluorescence from protein-DNA crystals

Our aim is to correlate 2-aminopurine fluorescence changes with structural changes in protein-DNA complexes. Our objectives are:

a) to solve, using X-ray crystallography, the atomic structure of the *Hha*I methyltransferase in complex with DNA containing 2-aminopurine in an extra-helical "base-flipped" position and in a second crystal form with 2-aminopurine in a normal base-paired conformation.

b) to record and analyse the fluorescence properties of the 2-aminopurine in both of these locations in the crystals using picosecond time-resolved fluorescence spectroscopy (time-correlated single photon counting method, TCSPC).

Crystals structures of complexes of *Hha*I bound to DNA have been solved for a wild type protein and for a mutant protein. The 2-aminopurine (2AP) was incorporated at the methylation site and clearly showed that it was flipped out of the DNA by the enzyme. The 2AP conformation was slightly different in the two complexes.

Prior to performing detailed time-correlated single photon counting (TCSPC) fluorescence analysis of the crystals, the behaviour of the 2AP fluorophore in aqueous solution needs to be better understood. A complete set of data on the behaviour of DNA containing 2AP at various locations in the DNA sequence has been obtained in aqueous solution. The changes observed upon addition of the *Hhal*

enzyme have also been fully characterised. The data obtained from a TCSPC instrument (in collaboration with Dr. David T.F. Dryden, and Prof. Anita Jones, University of Edinburgh, UK) show that the 2AP fluorescence is exquisitely sensitive to its environment, with only the base-flipping methylation location showing any changes in TCSPC signal. Differences in 2AP behaviour at this base-flipping site are observed between a wild type protein and a mutant and these changes correlate with a slight difference between the crystal structures determined. Furthermore, these differences also correlate with the behaviour of 2AP alone in solution, as our measurements are showing novel evidence of tautomerism and solvent effects.

Preliminary TCSPC data on crystals of *Hha*I bound to DNA containing 2AP at the methylation site have been obtained and show that the crystal constrains the 2AP to adopt only the base-flipped position as observed from the X-ray structure. The behaviour of 2AP at other locations still needs to be studied but we expect the crystal to constrain the 2AP to remain within the DNA helix and alter the TCSPC behaviour. Such a constraint would be expected to remove the TCSPC component due to extrahelical motion of the base which naturally occurs in solution and which is accentuated by base-flipping enzymes such as *Hha*I (manuscript in preparation).

Structure-function analysis of mammalian DNA methyltransferases

In higher eukaryotes including mammals, DNA methylation plays a key role in the control of numerous cellular processes such as transcription, genomic imprinting, developmental regulation. Three major types of MTases are known in mammals, which are large multidomain proteins involved in intricate interactions with other cellular components. We study the mechanism and interactions of DNMT1 and DNMT3a,b MTases using mutagenesis, chemical foot-printing, kinetic analysis and fluorescence spectroscopy.

The murine MTases Dnmt 1, Dnmt 3a and Dnmt 3b were isolated from baculovirus infected SF9 cells as homogeneous full-length proteins (in collaboration with Prof. Shoji Tajima, Institute for Protein Res., Osaka University, Japan). The MTases were interrogated using a series of synthetic DNA duplexes containing a unique hemimethylated CpG site in which the fluorescent base analog 2-aminopurine (2AP) has been introduced in different positions. 2AP fluorescence titrations and catalytic activity studies revealed that the exocyclic O6 atom of the 3' adjacent guanine on the target strand is essential for Dnmt1 activity, while the 2-amino group is not. More importantly, we show that while M.*Hha*I rotates only the target nucleotide on one strand out of the DNA helix during catalysis, Dnmt1 exhibited behavior that is consistent with unstacking of the bases on both strands of the target site. Real-time direct observation of the flipping motion using the stopped-flow technique indicates that the open reaction intermediate is assembled at a rate of ~0.4 s^{-1} at 25°C, which is faster than the observed multiple turnover rate of Dnmt1. In contrast, no detectable interaction with the hemimethylated substrate was observed in the case of Dnmt3a and Dnmt3b. In aggregate, our study confirms the expected general mechanistic similarity of all cytosine-5 MTases, while the observed differences in how MTases interact with DNA correlate with their distinct functional roles *in vivo*.

Methyltransferases as tools for covalent modification of DNA

One of the ultimate goals of this study is to develop new methods for sequence-specific covalent modification of cytosine residues in DNA. Using the cytosine-5 DNA MTase M.*Hha*I as a model system, we are exploring two approaches (in collaboration with Prof. Elmar Weinhold, Institute of Organic Chemistry, RWTH Aachen, Germany). The nucleobases in double-stranded DNA and stacked in the interior of the molecule and are hardly accessible to chemical modifications. Nucleotide flipping by MTases could in principle be used to sterically activate individual nucleobases for reactions with exogenous chemical reagents. We are designing variants of *Hha*I MTase with enhanced accessibility towards the flipped out cytosine and searching for chemicals that would specifically attach a suitable anchor group to DNA. To achieve direct chemical modification of flipped out target bases, we engineered M.*Hha*I for accessibility of the target cytosine. Inspection of molecular models for ternary M.*Hha*I-DNA-cofactor complexes suggests that the target base is obstructed by the catalytic loop and is hardly accessible. Two variants of M.*Hha*I, containing large deletions in the catalytic loop were constructed and purified and characterized. They showed good solubility and high base flipping potential, and will be used in further experiments.

Another strategy to achieve sequence-specific modification of DNA is to redesign the catalytic function of a MTase such that flipped out bases would be processed in a manner that is more useful than methylation. For example, MTases strongly bind DNA containing mismatches or cause deamination of cytosine in DNA when cofactor AdoMet is absent. We use rational design and molecular evolution approaches to produce variants of M.*Hha*I that would excise or deaminate flipped out cytosine residues. Production of a sequence-specific cytosine DNA glycosylase by redesign of the catalytic domain of M.*Hha*I is another line of research. We started rational redesign the catalytic site by computer-guided (collaboration with Dr. Janusz Bujnicki, International Institute of Cell and Molecular Biology, Warsaw, Poland) introduction of directed mutations at positions R163, R165 and T250. Altogether, 30 single mutants of M.*Hha*I were constructed, purified and largely characterized. Based on their catalytic properties, best candidates are being selected for further studies to produce combinations of double and triple mutants.

N4,5-dimethylcytosine, a novel hypermodified base in DNA

Three types of enzymatic methylation products are known in DNA: N4-methylcytosine, 5-methylcytosine or N6-methyladenine. Modification of DNA by two different MTases recognizing identical or overlapping sequences and normally targeting either 5- or N4-position of the same cytosine residue, can lead to the formation of a doubly methylated base - N4,5-dimethylcytosine (mmC). For example, such double methylation of cytosine (underlined) at the CCWGG sites *in vitro* is achieved when the C5-MTase *Eco*dcm acts first followed by the N4-specific *Mva*I enzyme. The formation of mmC *in vivo* and its consequences for the viability of *E. coli* was studied in *dcm*+ strains carrying the *Mva*I MTase gene on a plasmid. We previously found that the transformation of the plasmids was dramatically restricted in the mutants that are deficient in DNA repair, or leads to the induction of the *E. coli* SOS system. Our current biochemical analysis of cytosine modification in DNA at different expression levels of the *Mva*I MTase showed that both monomethylated cytosines were present at varied ratios; mmC was barely detectable by conventional chromatographic analysis of nucleosides, but was identified using a coupled ESI-MS detector. Altogether, these observations indicate that mmC residues do form *in vivo*, but are effectively removed from DNA. The structural basis for the observed low tolerance of mmC residues in DNA derives from molecular modelling, which revealed that mmC introduces local perturbations in the double helical DNA. Moreover, the action of certain DNA polymerases (Taq, Klenow, Sequenase) *in vitro* was significantly interrupted at sites of doubly-methylated cytosines. These findings suggest that enzymatically generated mmCs can be used as site-specific structural probes for studying the mechanisms of DNA damage recognition and repair *in vivo*.

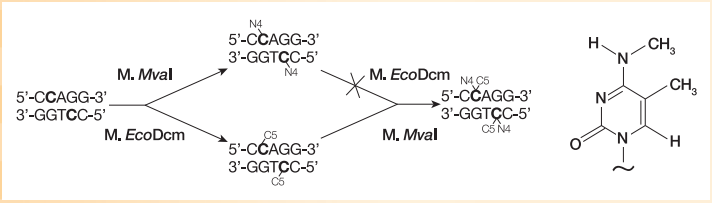


Fig. 4. Enzymatic formation of N4,5-dimethylcytosine (right) in DNA by successive action of the *Eco*dcm and *Mva*I methyltransferases.

Collaboration

Prof. Dr. Thomas Szyperski, State University of NY, Buffalo, USA
Dr. Christoph W. Mueller, EMBL Grenoble, France
Dr. David T. Dryden, Edinburgh University, UK
Prof. Dr. Xiaodong Cheng, Emory University, Atlanta, USA
Prof. Dr. Elmar Weinhold, Institute of Organic Chemistry, RWTH Aachen, Germany
Prof. Shoji Tajima, Institute for Protein Research, Osaka University, Japan
Dr. Peter Varnai and Dr. Richard Lavery, IBPC, Paris, France
Dr. Janusz Bujnicki, International Institute of Molecular and Cellular Biology, Warsaw, Poland

Grants

Howard Hughes Medical Institute, USA: International Research Scholars Programme
Volkswagen-Stiftung, Germany: Cooperation with Scientists in Central Europe
NATO Science Programme: Collaborative Linkage Grant
Royal Society: Joint Projects under European Science Exchange Programme
DAAD, Germany: International Quality Network
Lithuanian State Science and Studies Foundation

Recent publications

Vilkaitis G., Lubys A., Merkienė E., Timinskas A., Janulaitis A. & Klimašauskas S. (2002). Circular permutation of DNA cytosine-N4 methyltransferases: *in vivo* coexistence in the *Bcn*I system and *in vitro* probing by hybrid formation. ***Nucl. Acids. Res.*** 30:1547-1557.

Klimašauskas S., Gerasimaitė R., Vilkaitis G., & Kulakauskas S. (2002). N4,5-dimethylcytosine - a novel hypermodified base in DNA. ***Nucl. Acids Res. Suppl.*** (2):73-74.

Daujotytė D., Vilkaitis G., Manelytė L., Skalicky J., Szyperski T., & Klimašauskas S. (2003). Solubility engineering of the *Hha*I methyltransferase. ***Prot. Engineer.*** 16:295-301.

Daujotytė D., Serva S., Vilkaitis G., Merkienė E., Venclovas Č., & Klimašauskas S. Gln237 plays a key role in active flipping of the target cytosine by the *Hha*I methyltransferase. ***Structure*** 2004, submitted.

Contact: klimasau@ibt.lt

Laboratory of Prokaryote Genetic Engineering

Head Dr. Arvydas Lubys

Kornelijus Stankevičius, Ph.D.

Albertas Timinskas, Ph.D.

Mindaugas Margelevičius, Ph.D.

Renata Rimšelienė, M.Sc.

Arturas Jakubauskas, M.Sc.

Sonata Jurenaitė-Urbanavičienė, M.Sc.

Edita Kriukienė, M.Sc.

Rimantas Šapranauskas, M.Sc.

Elena Armalytė, M.Sc.

Jolanta Giedrienė, M.Sc.

Eglė Rudokienė, M.Sc.

Mantas Žvironas, B.Sc.

Jurgita Šerkšnaitė, B.Sc.

Linas Mažutis, student

Kristina Jatkovska

The research objectives of the Laboratory of Prokaryote Genetic Engineering are focused on the investigation of different aspects of the restriction-modification (RM) phenomenon. Restriction-modification systems function to protect the host cell from the invasion of foreign DNA. They are classified into three main types, according to their subunit composition, cofactor requirements and some other features. All RM systems comprise a pair of enzymatic activities - a DNA-methyltransferase (MTase) and a restriction endonuclease (REase), both recognizing the same short (4-8 bp) nucleotide sequence. Restriction endonucleases catalyze double-stranded cleavage of DNA, while the function of methylases is to protect hosts' genomic DNA from cleavage by cognate specific endonucleases.

Currently, we investigate:

- mechanisms by which RM systems are regulated
- type II bifunctional restriction endonucleases
- type IIT restriction endonucleases
- type IIS restriction endonucleases

In parallel, we are introducing the technique of microarrays, which we plan to use in parallel analysis of expression of some *Helicobacter pylori* genes. Also, we are trying to improve existing algorithms in order to get higher sensitivity of comparative analysis. Finally, in 2003 we started the medium-scale sequencing project in order to acquire knowledge about techniques and programs necessary in similar projects.

Regulation of restriction-modification systems

The two genes of RM system are always located in the close proximity to each other in bacterial chromosomes or plasmids. On many occasions complete RM systems were cloned in one step. This observation implies that some mechanisms regulating coordinated appearance of modification and restriction activities operate during the entry of recombinant plasmids with RM genes into a new host. How restriction-modification systems are regulated? The expression of genes in several RM systems has been shown to be regulated by small C proteins which contain helix-turn-helix motif. The analysis of *Esp*1396I RM system cloned and sequenced in the Laboratory revealed the presence of small gene that also codes for the regulatory protein. Fusion techniques were used to demonstrate that *Esp*1396I RM system uses the mechanism of coupled translation to ensure more tight control of the restriction endonuclease gene (Cesnaviciene et al., 2003). Currently, we are finishing the investigation of unique regulation of *Eco*255I RM system.

Organization of unusual RM systems

Investigation of structural organization of RM systems is carried out mainly throughout the cloning of restriction-modification genes, determination of their nucleotide sequences and comparison of cloned genes and encoded proteins with other restriction-modification systems. More than 70 complete R-M systems have been cloned and approximately half of them sequenced in the Laboratory. Recently, we have demonstrated that *Alo*I RM system, which recognizes the sequence $\downarrow$7/12-13(N)GAAC$(N)_6$TCC(N)12-13/7$\downarrow$ and cleaves both DNA strands on both sides at positions indicated by arrows, comprises only one large protein which *in vivo* and *in vitro* performs both restriction and modification. Basing on comparative analysis we have suggested that the N-terminal region of AloI is responsible for the DNA cleavage, the central part - for DNA methylation, and C-terminus - for DNA target recognition. In 2002, we cloned and sequenced the *Ppi*I RM system which recognizes a DNA target of very similar structure: $\downarrow$7/12(N)GAAC$(N)_5$CTC(N)13/8$\downarrow$. We observed that structural organization of *Ppi*I resembles that of *Alo*I. In 2003, we performed experiments aimed to elucidate function of aforementioned regions experimentally. First, the putative catalytic site of both *Alo*I and *Ppi*I was mutated and catalytic properties of resulting mutants investigated. We observed that mutations eliminated completely the ability of mutants to cleave DNA, however, their DNA methylation properties remained unaltered. This finding supports our theoretical predictions and indicates that DNA methylation function does not depend on the endonucleolytic function. Second, we performed domain swapping experiments between *Alo*I and *Ppi*I and determined the activity and specificity of resulting hybrids. The results clearly indicated that C-terminal parts of enzymes are interchangeable and that they are responsible for the target recognition. Moreover, we found that highly similar *Alo*I and *Ppi*I regions, which we predicted to be responsible for the recognition of tetranucleotide GAAC, are interchangeable as well and their

replacement results in mutants of wild-type specificity. In contrast, exchange of protein regions responsible for the recognition of trinucleotide components of complex targets resulted in specificity of hybrids which was the same as the specificity of enzyme which have donated corresponding protein region. This finding allows to suggest that the distance between tetra- and trinucleotide components of *Alo*I and *Ppi*I targets is defined by the protein region that is responsible for the recognition of trinucleotide sequences (manuscript in preparation).

Investigation of type II bifunctional restriction endonucleases

Type II restriction endonucleases specifically recognize and cleave DNA. Due to this property, they have become widely used tools in molecular biology, biotechnology and diagnostics. However, their usefulness depends on the range of specificities available. Demand for the new specificities was one of the driving forces that stimulated elucidation of molecular mechanisms involved in DNA recognition and cleavage, expecting to generate new REase specificities experimentally. We have developed a new procedure called the Methylation Activity Based Selection (MABS) for generating REases with a new specificity. MABS uses a unique property of bifunctional type II REases to methylate DNA targets they recognize. The efficiency of this technique was demonstrated by altering target specificity of the bifunctional REase Eco57I. This enzyme cleaves DNA 16/14 nucleotides away from the asymmetric recognition sequence CTGAAG and modifies the second A base within the upper strand of the target duplex (underlined). *Eco*57I mutants with altered substrate specificity CTGRAG were isolated. It was the first successful example of experimental generation of the type II REase with a new specificity (Rimseliene et al., 2003; patent applications EP1179596A1 and US 20030040614). In parallel experiments, we explored this technique in order to select for *Alo*I mutants of altered specificity.

Investigation of type IIT restriction endonucleases

Type IIT restriction endonucleases recognize asymmetric DNA targets and cleave DNA (at least one DNA strand) within the target. It is well known that DNA recognition and catalysis are tightly coupled in orthodox type II enzymes (they recognize and cleave symmetric DNA targets and act as homodimers). Unfortunately, there are no studies analyzing the mode of recognition and catalysis of type IIT enzymes, and LPGE scientists were first who cloned the system representing this scarce group (*Bpu*10I RM system). We observed that *Bpu*10I restriction endonuclease is composed of two non-identical subunits. Later, we demonstrated that each subunit contains an active site, which is responsible for the cleavage of only one DNA strand, and that inactivation of each of them results in creation of site- and strand-specific nickase (patent application EP1176204). In order to test if such organization is characteristic feature of this group of enzymes, we cloned and sequenced *Mva*1269I RM system. The unusually large amino acid sequence of restriction endonuclease *Mva*1269I and properties of mutants of its putative catalytic motif allowed us to suppose that *Mva*1269I possesses two active centers in one polypeptide chain. Catalytic mutants displayed chain-specific nicking activity on asymmetric DNA sequence recognized by wt *Mva*1269I. This differs sharply from results of inactivation of catalytic center in few other restriction endonucleases where a complete loss of endonucleolytic activity was observed. Also, *Mva*1269I was found to be in monomeric form when free or bound to specific DNA. Further investigations showed interdependent cleavage of two strands. This work revealed facts supporting the hypothesis that monomeric *Mva*1269I has two catalytic centers each of which is responsible for the cleavage of separate DNA strands (manuscript in preparation).

Investigation of type IIS restriction endonucleases

Type IIS restriction endonucleases recognize asymmetric targets and cleave both DNA strands outside the target. Only one enzyme of this group, *Fok*I, was analyzed in details. Proteolytic and X-ray studies have indicated that *Fok*I is composed of two domains, one of them being responsible for DNA recognition and the other one - for cleavage. In order to test if other representatives of this group have similar structural organization, we cloned, sequenced and investigated type IIS restriction-modification systems *Bfi*I and *Mnl*I.

The *Bfi*I domain organization was investigated using a new protein fragmentation technique, which was developed in our Laboratory and called Random Insertional Mutagenesis (RIM). In fact, RIM mimics reverse evolution, the process in which a two-domain protein

is dissected into two individual domains. The technique is based on a random DNase I / S1 - mediated disruption of a gene under investigation followed by insertion of a specially designed cassette. The cassette ensures premature termination of N-terminal protein fragment and independent translation of C-terminal fragment if inserted in frame. Application of RIM technique allowed us to determine exact boundaries of structural/functional domains in *Bfi*I and isolate mutants which retain their wild-type specificity but are composed of two complementing oligopeptides. We demonstrated that *Bfi*I endonucleolytic domain is N-terminally located, whereas the C-terminal domain performs target recognition. The order of domains in *Bfi*I is opposite compared to that of *FokI* (manuscript in preparation).

The *Mnl*I restriction endonuclease *Mnl*I recognizes CCTC targets and cleaves upper DNA strand 7 nt and bottom strand 6 nt to the right. The C-terminal part of cloned and sequenced *Mnl*I restriction endonuclease revealed intriguing similarity with the nuclease domain of "H-N-H" type of bacterial toxins, Colicin E7 and Colicin E9, and allowed us to identify the conserved sequence motif $Rx_3ExHHx_{14}Nx_8H$ of *Mnl*I as a new putative catalytic site in the family of restriction enzymes. A set of alanine replacements in putative active site was constructed and investigated. We observed that mutations of conserved amino acids have only minor effect on the specific DNA binding but can cause sharp reduction or even total inactivation of catalytic activity. Moreover, *Mnl*I cleaves both double stranded and single stranded DNA in the presence of magnesium and nickel ions, respectively, like above mentioned Colicines. These results indicate that *Mnl*I differs from *FokI*, the member of so-called "PD-EXK" superfamily, and *Bfi*I, the recently described member of phospholipase D superfamily, by both the active site architecture and the reaction mechanism, and in fact represents a new subtype of type IIS restriction endonucleases (manuscript in preparation).

In an ongoing search for evolutionary related REases that recognise partially overlapping target sites, we have cloned and sequenced genes of three type IIS RM systems, *Alw*26I (GTCTC), *Eco*31I (GGTCTC) and *Esp*3I (CGTCTC), which target sites share the common pentanucleotide, GTCTC. Our study showed that the deduced aa sequences of the *Alw*26I, *Eco*31I and *Esp*3I REases share significant level of identity. In a pairwise comparison, the amino acid identity score is higher than that observed hitherto among the REases which recognise partially overlapping nucleotide sequences. With this information in hand, *Alw*26I, *Eco*31I and *Esp*3I REases represent an exceptionally attractive group of proteins to study structure-function relationships in DNA recognition. In addition, the accompanying methyltransferases of the respective restriction-modification systems, are found to be represented by proteins of an unusual structural and functional organisation (Bitinaite et al., 2002).

The role of the putative H. pylori species-specific DNA methyltransferase HpyI

Colonization of the gastric mucosa by *Helicobacter pylori* results in an acute inflammatory response and damage to the gastric epithelium. Inflammation can then progress to several disease states, ranging in severity from superficial gastritis, peptic ulceration, to mucosa-associated lymphoma and gastric cancer. An understanding of *H. pylori* pathogenesis, identification of virulence factors and their role in host-pathogen interactions is a prerequisite for developing effective measures for eradication of these bacteria. Our research is focused on the elucidation of the role of the putative *H.pylori* species-specific DNA methyltransferase *Hpy*I in expression of the genes which are or might be potentially responsible for pathogenesis of the bacteria.

Generation of H. pylori mutants deficient in species-specific MTase. The sequencing of *H.pylori* genome revealed an extraordinary high number of ORFs encoding DNA MTases. Bacterial MTases, which are taxono-specific (species-, family-specific, i.e. they are present in every strain of a given taxon) are involved in general cellular processes such as mismatch repair, regulation of DNA replication, DNA transposition or even may be indispensable for cell viability. The involvement of the orphan DNA adenine MTase (Dam) in regulation of pili phase variation in pyelonephritis associated *E. coli* is well documented. It has been shown that Dam MTase of *Salmonella typhimurium* regulates the expression of at least 20 virulence genes known to be induced during infection and the *dam*- mutants are totally avirulent. The homologue of the orphan Dam MTase is absent in *H. pylori*. Therefore its functional analogue, if it exists, remains to be identified in *H.pylori*. Our experiments on screening of more than 40 *H. pylori* strains for REases and MTases has led to the identification of a candidate species-specific *H.pylori* DNA methyltransferase. We found that CTAG targets recognised by *Hpy*I MTase are modified in all strains tested, implying that M.*Hpy*I is a good candidate for the *H. pylori* species-specific MTase. Therefore, we have deleted the gene for *Hpy*I methyltransferase in strain J99 in order to evaluate the effect of the mutation on the expression of genes known or expected to be involved in pathogenesis.

Development of H. pylori DNA array. The putative pathogenic determinants of *H. pylori* can be divided into two major groups: virulence factors, which contribute to the pathogenetic effects of the bacterium, and maintenance factors, which allow the bacterium to colonise and remain within the host. We have focused our research on the investigation of a number of genes responsible for motility and chemotaxis, utilisation of urea (including structural and accessory genes, urea and nickel transport, and chaperones), regulatory functions, transformation and cell killing, genes of *cagA* associated pathogenicity island (*cag*PAI), lipopolisacharides (LPS), outer membran proteins (OMP), adhesins, restriction-modification systems. Number of genes encoding enzymes, such as phospholipases, catalase, superoxide dismutase and some others were also included. Three genes encoding *Cfr*9I, *Eco*72I and *Mva*I restriction endonucleases and seven different fragments of bacteriophage λ DNA were also included as negative and spike controls, respectively. The primers for amplification of all selected genes were designed in such a way to yield PCR products of ~ 300 base pairs long which are prone to cross-hybridize with members of *H.pylori* paralogous gene families. They were designed in collaboration with Dr. B. Korn (Resource Center of the German Genome Project, Heidelberg, Germany). All PCR products were analysed for their integrity and absence of non-specific DNA bands.

Construction of DNA arrays. Prepared PCR fragmenst were spotted onto amino silane coated glass slides acceptable for GenePix 4000B DNA array scanner. Every single probe was gridded in quadruplicate. Controls for normalization were also printed. DNA array was printed in laboratory of Dr. J. P. Dumanski (Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden). We have scanned some slides using GenePix4000B DNA array scanner to assure that DNA was successfully spotted. All spots had similar shape and sufficient signal intensity. The same result was obtained after the staining of DNA microarray with SYBR Green II. Basing on these observations we concluded that the prepared DNA microarray is suitable for hybridization experiments.

Higher sensitivity comparisons

Comparison of DNA or protein sequences is the main step in solving key questions in life sciences such as structure/function of main active molecules or evolutionary relationship of living organisms. In some cases it is the only reliable way for the moment. There are numerous of different algorithms and software useful for similarity searches between sequences, the most popular and powerful to date are FASTA and BLAST and especially - a recent modification of the latter (PSI-BLAST). The main goal of our work was to investigate influence of further modifications of PSI-BLAST algorithm on software sensitivity to find similarities within protein sequences. Two parameters, namely, scan window size W and threshold score T, were varied. Test was carried out on DNA methyltransferases (MT) previously described to be homologous. Two groups, i.e. alpha group of amino MT (possessing medium level similarities) and C5 MT (high similarities) were selected as model protein sets. As it was expected, extension of window size from common 3 to 4 or 5 resulted in a slight increase of PSI-BLAST sensitivity though at cost of calculation time. However, no positive effect on selectivity was obtained increasing threshold score.

Enjoying a challenging discussion with the famous biochemist Dr. Richard Roberts (right) from the New England BioLabs after the meeting at the Institute of Biotechnology in September 2003.

Grants

Lithuanian State Science and Studies Foundation

Contracts

UAB Fermentas, Lithuania, Vilnius

Recent publications

Vilkaitis G., Lubys A., Merkienė E., Timinskas A., Janulaitis A., Klimašauskas S. (2002). Circular permutation of DNA cytosine-N4 methyltransferases: *in vivo* coexistence in the *Bcn*I system and *in vitro* probing by hybrid formation. ***Nucleic Acids Res.*** 30:1547-1557.

Bitinaite J., Mitkaite G., Dauksaite V., Jakubauskas A., Timinskas A., Vaisvila R., Lubys A., Janulaitis A. (2002). Evolutionary relationship of *Alw*26I, *Eco*31I and *Esp*3I, restriction endonucleases that recognise overlapping sequences. ***Mol. Genet. Genomics*** 267:664-672.

Česnavičienė E., Mitkaitė G., Stankevičius K., Janulaitis K. and Lubys A. (2003). *Esp*1396I restriction-modification system: structural organization and mode of regulation. ***Nucleic Acids Res.*** 31:743-749.

Rimseliene R., Maneliene Z., Lubys A. and Janulaitis A. (2003). Engineering of restriction endonucleases: using methylation activity of the bifunctional endonuclease *Eco*57I to select the mutant with a novel sequence specificity. ***J. Mol. Biol.*** 327:383-391.

Patents

A.Janulaitis, K.Stankevicius, A.Lubys, A.Markauskas. Nuclease. EP1176204. 2002-01-30.

A.Janulaitis, R.Rimseliene, A.Lubys. Nuclease. EP1179596. 2002-02-13.

A.Janulaitis, R.Rimseliene, A.Lubys. Nuclease. US20030040614. 2003-02-27.

Contact: lubys@ibt.lt

Laboratory of Eukaryote Genetic Engineering

Head Dr. Habil. Kęstutis Sasnauskas

Alma Gedvilaitė, Ph.D.

Aušra Ražanskienė, Ph.D.

Danguolė Bartkevičiūtė, Ph.D.

Rimantas Šiekštelė, M.Sc.

Raimundas Ražanskas, M.Sc.

Aistė Bulavaitė, M.Sc.

Rasa Petraitytė, M.Sc.

Rimantas Slibinskas, M.Sc.

Kristina Vagonytė, B.Sc.

Eglė Aleksaitė, B.Sc.

Paulina Grigaitė, B.Sc

Mindaugas Juozapaitis, student

Helena Kondrotovič

Combined immune and gene therapy of chronic viral hepatitis

Many infectious agents persist in a patient because of an insufficient immune response. Natural viral capsids are highly immunogenic but often they do not induce neutralizing antibodies or a CTL response. Using genetic methods protective B and T cell epitopes can be introduced into capsid proteins and expressed in bacteria or yeast. It was planned *in vitro* to encapsidate such chimeric capsids into genes encoding cytokines, which favor Th1 and subsequent CTL responses. Recognition of such capsids by B cells led to efficient antibody production and Th1 cell epitope presentation in a favorable cytokine milieu. Suitability of such hepatitis B and polyoma virus capsid vectors were tested in cell cultures, and in mouse models.

For the construction of immunogenic chimeric VLPs forming proteins, polyomaviruses MPyV and SV40 VP1 proteins were used as carriers of selected HBV epitopes. Polyomavirus virus-like particles (VLPs) can be generated by heterologous expression and spontaneous self-assembly of major surface protein VP1. Although non-infectious, VLPs resemble virions in their strong immunogenicity and can be used as safe and efficient vaccines. VLPs tolerate the insertion and surface exposure of foreign epitopes and potentiate their immunogenicity.

First of all, both SV40 and MPyV VP1s were expressed in the yeast *Saccharomyces cerevisiae* in order to test the efficiency of VLPs formation in the yeast expression system. Results of expression confirmed, that VP1s of both viruses efficiently formed VLPs in yeast.

The yield of CsCl purified particles made up approximately 30-40mg/l liter of yeast culture.

For the construction of chimeric VLPs, four potential sites were selected in the structure of both viruses VP1 on the basis of three-dimensional structure of SV40 VP1. The selected sites were localized in the loops exposed to the surface of virus particle: Site 1 - BC-loop, 73-74 aa; Site 2 - EF-loop, 198-199 aa; Site 3 - F-G1-loop, 228-229 aa; Site 4 - HI-loop, 273-274 aa. The sites for the restriction endonuclease *Bgl*II were inserted in the selected position of the SV40 and MPyV VP1 genes by using specific primers and PCR (#4-Annexes I and IV).

For the construction of chimeric VLPs, three different HBV PreS and S sequences, encoding main B and T cell epitopes were selected and inserted into modified VP1 genes. Two PreS1 sequence variants encoding the most important PreS1 epitopes were selected. The first PreS1 sequence variant included PreS1-20-47aa fragment, the second variant included all PreS1 sequence except the hydrophobic part between 58 - 90 aa - 13-58//90-119 aa. The third variant was S sequence encoding main antigenic loop of the S protein (101-169aa), exposed to the surface of HBV particles.

HBV *ayw* serotipe, provided by partner #2 was used in these experiments.

All three HBV sequences were inserted into four different SV40 VP1and MPyV VP1 positions. The resulted hybrid genes were inserted into yeast expression vector and expressed in the yeast *S. cerevisiae*. Expression level and VLPs formation were investigated. Chimeric proteins that formed sufficient yield of VLPs (6 derived from SV40 VP1 and 4 derived from MPyV VP1) were selected for further investigations. Chimeric proteins containing two different inserts in the different VP1 loops of the same protein molecule were constructed also. CsCl and EM analysis of such chimeric proteins revealed, that both SV40 VP1 and MPyV VP1 tolerated double large insertions. The chimeric SV40 and MPyV VLPs containing HBV S sequence (101-169 aa) inserted in the BC loop and PreS1 sequence (13-59/90-119) inserted in HI loop were constructed and purified.

The purified chimeric VLPs were provided to consorcium partners (Dr. K.Schwarz, Zuerich) for the investigation of immunogenicity of chimeric particles.

The initial transfection of model gene (GFP) demonstrated, that chimeric particles as well as wt VP1 were not efficient tools for gene delivery, in comparison to transfection agents like Exgene (Fermentas UAB, Vilnius).

Hamster polyomavirus-derived virus-like particles as a promising universal carrier for various vaccine development applications

Protein engineering, i.e. generation of proteins with desired features, is a fascinating opportunity not only in producing of "optimized" food but also highly efficient and safe vaccines against different pathogens or for tumor therapy. Virus-like particles (VLPs) have been generated by heterologous expression of viral capsid and envelope proteins, which have an intrinsic capability to self-assemble into highly organized particles, often without the need of additional viral components. VLPs can induce a strong humoral immune response because of correct folding of the monomeric proteins, the resulting formation of conformational antigenic determinants, the multimeric structure of identical subunits, and the presence of TH epitopes. Moreover, VLPs are able to stimulate cytotoxic T cells. In addition to inducing strong immune responses VLPs are non-infectious and highly safe because of the lack of a viral genome and can therefore represent an ideal vaccine. VLP-based vaccines have been already in routine use for immunization against hepatitis B virus and tested for their efficiency to protect against papilloma- and rotavirus infections in animal models.

Epitopes and receptor-binding ligands from proteins of pathogens, which are not able to form VLPs, can be grafted onto VLPs by genetic engineering to improve their immunogenicity. The localization of potential insertion sites for foreign peptide sequences is essential for the generation of chimeric VLPs as well as the evaluation of foreign peptide size limits that can be inserted. In fact, fusion of longer or multiple copies of peptides/epitopes to VLPs usually introduces structural difficulties in protein folding and VLP formation.

The hamster polyomavirus (HaPyV) major capsid protein VP1 belongs to a family of polyomavirus VP1 proteins containing highly conserved structure motifs and functional domains. Yeast-expressed HaPyV-VP1 well tolerated the insertion of a foreign pentapeptide - an epitope with the sequence DPAFR originated from the preS1 region of HBV - at all four predicted surface-exposed sites between amino acid (aa) positions 80-89 (site #1), 222-225 (site #2), 243-247 (site #3) and 288-295 (site #4) of VP1. The insertion into VP1 did not influence protein folding and assembly of VLPs, and even allowed the generation of high-titered preS1-specific antibodies in mice. The high carrier competency of HaPyV-VP1 is also reflected by the fact that it tolerates not only single insertion of foreign sequences, but also the simultaneous insertion into different sites.

In this study we investigated the newly established carrier HaPyV-VP1 in comparison to other carrier systems, in terms of insertion capacity, antigenicity and immunogenicity. Therefore we inserted segments of *Puumala* hantavirus (strain *Vranica/Hällnäs*; PUUV-VR) nucleocapsid (N) protein containing 45, 80 or 120 N-terminal amino acid residues, previously used for insertion into HBV core-derived particles, into the four potential insertion sites of HaPyV-VP1. HaPyV-VP1 well tolerated the insertion of VR peptides at surface-exposed sites between aa positions 80 - 89 (site #1) or 288 - 295 (site #4) without influencing protein folding and assembly of VLPs. Site #4 of HaPyV-VP1 seems to tolerate even the insertion of green fluorescent protein. The insertion capacity at sites #2 (between aa 222-225) and #3 (between aa 243-247) is limited; insertions of short peptides are tolerated but longer foreign peptides derived from PUUV-VR N protein apparently affect correct folding and assembly of VLPs.

Therefore, in terms of insertion capacity, HaPyV-VP1-derived VLPs belong to the group of VLPs with a high insertion capacity like those tolerating more than 100 foreign aa as bluetongue virus NS1 tubules and parvovirus B19-derived VLPs or even more than 200 foreign aa as HBV core and Ty-derived VLPs. This insertion capacity is much larger than that observed for other VLP carriers, e.g. bacteriophage coat protein-derived VLPs.

Immunization with any of the chimeric VP1-derived VLPs containing VR inserts was able to induce a high-titered IgG antibody response against PUUV-VR N protein in BALB/c mice. The observed endpoint titres for PUUV-VR N protein-specific antibodies in mice immunized with chimeric VP1-derived VLPs harboring VR120-1 or VR120-4 inserts (about 1: 200,000) are in the same range as titers observed in BALB/c and C57/Bl6 mice immunized with HBV core particles harboring 120 N-terminal aa of N proteins from PUUV, Dobrava or Hantaan hantaviruses. Although the use of adjuvants induced about 10fold higher titers of PUUV-VR N-specific antibodies in mice immunized with VP1/VR120-1 and VP1-VR120-4, immunization without adjuvants also provided long-term humoral immune response against PUUV-VR N protein and induction of antigen-specific T-cell help.

The detection of N-specific IgG antibodies of both IgG1 and IgG2a subclasses in mice immunized with chimeric VP1/VR VLPs suggest a mixed Th1/Th2 response. Direct measurement of antigen-induced cytokine secretion by spleen cells *in vitro* is in line with the activation of antigen-induced T cells from both Th1 and Th2 subsets. The mixed IgG1/G2a response suggests an approximate Th1-Th2 cell balance in the regulation of Ig isotype switching.

We observed that the insertion of N protein segments reduced the VP1 carrier antigenicity and immunogenicity. The level of reduction is related to the size of the foreign insertion; i.e. the larger the insert the lower the remaining VP1 antigenicity and immunogenicity. This might represent an additional advantage of chimeric HaPyV-derived VLPs in terms of vaccine development where a low antigenicity and immunogenicity of the carrier itself is wanted.

In conclusion, HaPyV-VP1-derived VLPs were shown to be able to tolerate foreign protein segments of up to 120 aa at certain sites of VP1. These chimeric VLPs are able to induce a strong insert-specific antibody response, even in the absence of any adjuvant. Therefore, HaPyV-VP1 represents a promising carrier for future human tumor vaccine development based on HaPyV-derived VLPs presenting tumor antigen CTL epitopes. Various strategies have been evaluated concerning the best way of delivering T-cell epitopes to induce a tumor epitope-specific cytotoxic T-cell response. We suggest HaPyV-derived VLPs will tolerate the insertion of a human CTL tumor epitope from MUC1 (human mucin 1) at different sites of a single VP1 molecule and will investigate the capacity of chimeric VLPs to induce a peptide-specific antibody response in mice, the uptake of VLPs and maturation of human monocyte-derived DCs and the capacity of DCs to induce an efficient epitope-specific CTL response *in vitro*.

Hantavirus vaccine for Europe

Hantaviruses represent a separate genus *Hantavirus* of the family *Bunyaviridae*. They contain a tripartite RNA genome of negative polarity. The genome segments M (medium) and L (large) encode a glycoprotein precursor, which is co-translationally processed into G1 and G2, and an RNA-dependent RNA polymerase. The S (small) genome segment codes for the nucleocapsid (N) protein. Recently, we have described the high-level yeast expression of authentic and amino-terminally His-tagged rN protein of PUUV (strain *Vranica/Hällnäs*; most likely originated from Sweden), which are able to induce a protective immune response in bank voles, the natural host of PUUV (Dargeviciute et al., 2002).

High-level expression in yeast and purification of high yields of rN proteins of hantavirus species originating from Asia (HTNV) and different regions of Europe (PUUV strains Sotkamo and Kazan from Finland and Russia, respectively; DOBV-Slk from Slovakia and DOBV-Slo from Slovenia) was also performed. The rN proteins were characterized in terms of stability, nucleic acid and endotoxin contamination, antigenicity and immunogenicity.

Hantavirus N can be expressed at high level in yeast.

The yields of purified rN proteins from different hantaviruses mentioned above were reproducibly high ranging from 0.5 to 1.5 mg per g wet weight of yeast cells.

The main portion of rN protein in yeast remains insoluble and forms aggregates.

In line with observations during the expression of hantavirus rN proteins in *E. coli*, the yeast-expressed rN proteins of PUUV-Vra and the other hantaviruses described here were highly insoluble, requiring 8M urea for solubilisation. Nevertheless, as evidenced previously already for PUUV-Vra rN protein (Dargeviciute et al., 2002), the rN proteins are highly reactive with mAbs in Western blots. Moreover, rN proteins of PUUV-Vra, HTNV-Foj and DOBV-Slk have been used to establish highly specific and sensitive IgG ELISAs for the detection of antibodies in human sera.

Yeast-expressed rN proteins are highly pure and stable during long-term storage.

In terms of vaccine development such properties as protein stability, nucleic acid contamination and endotoxin contamination are worth to investigate. The yeast-expressed rN proteins are stable in yeast; when analysing yeast lysates obtained after induction of heterologous protein synthesis we did not observe a significant degradation of rN proteins in Western blots.

The nucleic acid contamination of PUUV-Vra rN protein determined by a dot-blot method was found to be very low (about 6.4 pg per 100 µg). This level of DNA contamination is much lower than the contamination accepted for commercial vaccines for human use of 100 pg per dose (see Center for Biologics Evaluation and Research and FDA, 1993). As expected, we found no endotoxin contamination in the preparation of purified PUUV-Vra rN protein.

Yeast-expressed rN proteins induce a strong and cross-reactive antibody response in rabbits.

In line with immunisation data in BALB/c and C57BL/6 mice the yeast-expressed rN proteins induced a high-titered, cross-reactive antibody response in rabbits reaching reciprocal end point titres of up to 10^5 to 10^6 for the corresponding homologous antigen. This strong immunogenicity is also in line with immunisation data in mice and bank voles using *E. coli*-expressed rN proteins of different hantaviruses.

In conclusion, our yeast expression system is very efficient for the generation of rN proteins of different hantaviruses originated from Asia and Europe. An amino-terminal His-tagging enables an easy purification of hantavirus rN proteins. The His-tagged rN proteins were found to be highly stable during long term storage (lyophilised at 4°C and -20°C, dialysed and stored at -20°C). The PUUV-Vra rN protein was demonstrated to be free of endotoxin contamination and to contain only very small traces of nucleic acid contamination. Moreover, immunisations of rabbits confirmed the high immunogenicity of the yeast-expressed rN proteins of the different hantaviruses. Therefore, non-infectious, yeast-expressed hantavirus rN proteins represent promising vaccine candidates. In addition, these rN proteins are helpful tools for establishing specific and reliable assays for hantavirus diagnostics.

Enhanced surveillance of measles

Measles virus, belonging to the *Morbillivirus* genus of a family *Paramyxoviridea*, is an enveloped virus, which encapsidates a 16-kb negative strand RNA genome (for review, Griffin, 2001). Mature particles are assembled from six viral-encoded proteins. Two glycoproteins, the hemagglutinin and the fusion protein, are the major protein constituents of the envelope. A third protein, a matrix protein, lines the interior surface of the viral envelope and is believed to link glycoproteins to the internal viral proteins. The nucleoprotein (MeN), the major internal protein, appeared to be responsible for viral genome packaging and formation of replication complexes along with other proteins. Two additional proteins, namely, the phosphoprotein and large protein or RNA polymerase, are associated with nucleocapsid and are involved both in replication and transcription of viral RNA. Viral nucleocapsid proteins usually elicit a strong long-term humoral immune response in patients and experimental animals. Therefore, measles diagnosis can be based upon the detection of antibodies specific to MeN. The main task for the Institute of Biotechnology, the Laboratory of Eukaryote Genetic Engineering was to develop a simple, efficient and cost-effective system for generation of measles virus nucleoprotein for measles serology.

The gene encoding measles virus *Schwarz* strain nucleoprotein was successfully expressed in two different yeast genera, *Pichia pastoris* and *Saccharomyces cerevisiae*, respectively. Both yeast genera synthesized a high level of nucleoprotein, up to 29% and 18% of total cell protein, in *P. pastoris* and *S. cerevisiae*, respectively. This protein is one of most abundantly expressed in yeast. Convalescent sera from measles patients reacted with the recombinant protein and a simple diagnostic assay could be designed on this basis. After purification nucleocapsid-like particles derived from both yeast genera appeared to be similar to those detected in mammalian cells infected with measles virus. A spontaneous assembly of nucleoprotein into nucleocapsid-like particles in the absence of the viral leader RNA or viral proteins has been shown. Compartmentalization of recombinant protein into large compact inclusions in the cytoplasm of yeast *S. cerevisiae* by GFP fusion has been demonstrated.

The purified rMeNP and a monoclonal antibody to the rMeNP conjugated to horseradish peroxidase were used to develop a measles specific IgM capture EIA, MACEIA, in both serum and oral fluid specimens. Evaluations of the MACEIA were performed by testing a) serum samples (n=80), b) paired oral fluid/serum samples (n=37, representing 15 cases) and c) oral fluid samples from cases of measles notified by general practicioners (n=145). The samples were also tested by measles IgM radioimmunoassay, MACRIA. The sensitivity and specificity of the MACEIA compared to MACRIA for a) the serum samples were 100% and 96.6%, respectively, b) for paired serum / oral fluids samples 100% and 100%, respectively and c) for oral fluid samples 96.6% and 100%, respectively.

Conclusion: Recombinant yeast derived measles nucleocapsid protein efficiently formed nucleocapsid-like particles in both yeast genera, *S. cerevisiae* and *P. pastoris*. Yeast generated MeNP nucleocapsid-like particles are suitable for the development of very sensitive measles specific IgM and IgG detection kits.

Investigation of protein secretion in yeast Kluyveromyces

Kluyveromyces lactis is biotechnologically significant yeast, which has been already exploited as a host for the production of heterologous proteins due to its secretory performance. Production of a foreign protein via the eukaryotic secretory pathway ensures high fidelity of folding, assembly and modification processes required for biological activity. Although high-level transcription of genes is attainable in eukaryotic systems, secretion of corresponding protein products often does not increase proportionally. In many cases, retention within the endoplasmic reticulum (ER) or the point of exit from the ER lumen into Golgi appears to be responsible for this bottleneck. Once within the ER lumen, a secretory protein must be correctly folded and in some cases, modified, assembled into a functional oligomer and released through the secretory pathway. The major rate-limiting step in constitutive protein secretion in eukaryotic cells occurs at the point of exit from the ER lumen into Golgi, and it is at this point quality control is exerted. Extensive cellular machinery responsible for efficient and accurate maintenance of these processes within endogenous cellular proteins exists, but if a cell is genetically modified to express a high level of a heterologous secretory protein, one or more of these processes may become rate or yield limiting.

Several publications have provided the approach to overcome such problems in *S. cerevisiae* by modulating cellular levels of soluble ER components, namely, protein disulphide isomerase, which catalyses the formation of native disulphide bonds in secretory proteins or BiP, which appears to act, in part, as a molecular chaperone for secretory proteins passing through the ER. Over-expression of syntaxins Sso1 and Sso2, functioning at the targeting/fusion of the Golgi-derived secretory vesicles to the plasma membrane, polyubiquitin, *PSE1* and *YAL0*48c ORF or the disruption of protease Yps1p, Yap3p encoding genes enhance heterologous protein secretion also. Quality and quantity of a product might be improved in mutants with an altered glycosylation process to prevent over-glycosylation or with a reduced vacuolar protease content to prevent proteolytic degradation. Recently it has been shown, that an increased dosage of polyubiquitin and *PDI1* genes enhanced human protein secretion in *K. lactis*.

An empirical approach might be fruitful in search for genes, which increase the rate and yield of secretory proteins, if over- or under- expressed. First of all it was demonstrated for the prochymosin producing strain with two recessive *ssc1* and *ssc2* mutations, in which the distribution of product between the vacuole and the medium was shifted towards an improved release. Later the *SSC1* gene was shown to be identical to the *PMR1* gene encoding Ca^{2+}-ATPase involved in the secretory pathway. Other mutations like *rgl2* act on transcriptional level.

Previously we isolated mutants responsible for a super-secretion phenotype in *K. lactis* using the gene coding for a *Bacillus amyloliquefaciens* α-amylase as a marker for secretion. The recessive mutant strain, which secreted the heterologous protein in the five-fold excess compared to the wild-type strain, was used for cloning of genes, restraining super-secreting phenotype in high copy number state. In screening for genes affecting super-secreting phenotype, we found, that the *KlMNN10* gene in multicopy state suppressed the super-secreting phenotype. We demonstrated, that the disruption of the *KlMNN10* in *K. lactis* as well as *MNN10* and *MNN11* in *S. cerevisiae* conferred the super-secreting phenotype. *MNN10* isolated from *S. cerevisiae* suppressed the super-secretion phenotype in *Kl. lactis klmnn10* strain, likewise homologous *KlMNN10*. The genes *MNN10* and *MNN11* of *S. cerevisiae* encode mannosyltransferases responsible for the majority of the α-1,6-polymerizing activity of the mannosyltransferase complex. These data agree with standpoint that the structure of glycoproteins in a yeast cell wall strongly influences on the export of homologous and heterologous proteins. The set of genes, namely, the suppressors of over-secreting phenotype, could be attractive for further analysis of gene functions, over-secreting mechanisms and for construction of new strains containing reduced yeast genomes, optimized for heterologous protein secretion.

Collaboration

Prof W.H.Gerlich, Justus-Liebig-University of Giessen, Institute of Medical Virology, Giessen, Germany
Prof. P.Pumpens, Biomedical Research and Study Center University of Latvia, Riga, Latvia
Dr. K. Schwarz, CYTOS Biotechnology AG, Zuerich, Switzerland
Prof. D.H. Krüger and Dr. R.Ulrich, Institute of Virology, Humboldt University, Berlin, Germany
Prof. Å.Lundkvist, Microbiology and Tumour Biology Center Karolinska Institutet, Stockholm, Sweden
Dr. D.Samuel and Dr. Jin Li, HPA, Enteric and Respiratory Virus Laboratory, London, UK
Dr. D.Becher, Micromun GmbH, Greifswald, Germany

Grants

EC Framework 5 programme, QLK2-CT-1999-01119.
EC Framework 5 programme, QLK2-CT-2000-01476.
EC Framework 5 programme, QLK2-CT-2001-01632.

Dr. Aušra Ražanskienė (left), Dr. Alma Gedvilaitė and Dr. Habil. Kęstutis Sasnauskas, winners of the Lithuanian Prize in Science of 2003 for *Synthesis of Virus Proteins in Yeast and their Application for Vaccines and Diagnostics*

Recent Publications

Sasnauskas K., Bulavaite A., Hale A., Jin L., Gedvilaite A., Dargevičiūtė A., Bartkevičiūtė D., Žvirblienė A., Staniulis J., Brown D. W. G., Ulrich R. (2002). Generation of recombinant virus-like particles of human and non-human polyomaviruses in yeast *Saccharomyces cerevisiae*. ***Intervirology*** 45 (4-6):308-317.

Pumpens P., Razanskas R., Pushko P., Renhof R., Gusars I., Skrastina D., Ose V., Borisova G., Sominskaya I., Petrovskis I., Jansons J., Sasnauskas K. (2002). Evaluation of HBs, HBc, and frCP Virus-Like Particles for Expression of Human Papillomavirus 16 E7 Oncoprotein Epitopes. ***Intervirology*** 45 (1):24-32.

Hale A., Bartkevičiūtė D., Dargevičiūtė A., Li Jin, Knowles W., Staniulis J., Brown D.W.G., Sasnauskas K. (2002). Expression and antigenic characterization of the major capsid proteins of human polyomaviruses BK and JC in *Saccharomyces cerevisiae*. ***J. Virol. Methods*** 104 (1):93-98.

Voronkova T., Grosch A., Kazaks A., Ose V., Skrastina D., Sasnauskas K., Jandrig B., Arnold W., Scherneck S., Pumpens P., and Ulrich R. (2002). The highly flexible and immunodominant C-terminal region of hamster polyomavirus VP1 exposed on chimeric bacteriophage fr capsids induced a strong VP1-specific antibody response in mice. ***Viral. Immunol.*** 15 (4):627-643.

Dargevičiute A., Brus S. K., Sasnauskas K., Kruger D. H., Meisel H., Ulrich R., and Lundkvist A. (2002). Yeast-expressed *Puumala* hantavirus nucleocapsid protein induces protection in a bank vole model. ***Vaccine*** 20:3523-3531.

Samuel D., Sasnauskas K., Li Jin, Beard S., Zvirblienė A., Gedvilaite A., Cohen B. (2002). High level expression of recombinant mumps nucleoprotein in *Saccharomyces* cerevisiae and its evaluation in mumps IgM serology. ***J. Med. Virology*** 66:123 – 130

Bartkevičiute D. and Sasnauskas K. (2003). Characterization of yeast *Kluyveromyces lactis mutations*, conferring super-secretion of recombinant proteins. ***Yeast*** 20 (1):1-11.

Stolt A, Sasnauskas K., Koskela P., Lehtinen M. and Dillner J. (2003). Seroepidemiology of the human polyomaviruses. ***J. Gen. Virol.*** 84:1499-1504.

Knowles W.A. and Sasnauskas K. (2003). Comparison of cell culture-grown JC virus (primary human fetal glial cells and the JCI cell line) and recombinant JCV VP1 as antigen for the detection of anti-JCV antibody by haemaglutination inhibition. ***J. Virol. Meth.*** 109:47-54.

Samuel D., Sasnauskas K., Li Jin, Gedvilaite A. Slibinskas R., Beard S., Zvirblienė A., Artimos S., Staniulis J., Cohen B., Brown D. (2003). Development of a measles specific IgM ELISA for use with serum and oral fluid using recombinant measles nucleoprotein produced in *Saccharomyces cervisiae*. ***J. Clin. Virol.*** 28:121-129.

Bartkevičiute D. and Sasnauskas K. (2003). Characterization of yeast *Kluyveromyces lactis* mutations, conferring super-secretion of recombinant proteins. ***Yeast*** 20 (1):1-11.

Slibinskas R., Samuel D., Gedvilaite A., Staniulis J., Sasnauskas K. (2004). Synthesis of the measles virus nucleoprotein in yeast Pichia pastoris and *Saccharomyces cerevisiae*. ***J. Biotechnol.*** 1072:115-124.

Gedvilaite A., Žvirblienė A., Staniulis J., Sasnauskas K., Krüger D. and Ulrich R. (2004). Segments of *Puumala* hantavirus nucleocapsid protein inserted into chimeric polyomavirus-derived virus-like particles induce a strong immune response in mice. ***Viral. Immunol.*** : 51-68.

Ražanskienė A., Geldmacher A., Schmidt J., Ritzi A., Niedrig M., Lundkvist Å., Krüger D., Meisel H., Sasnauskas K., and Ulrich R. High yields of stable and highly pure nucleocapsid proteins of different hantaviruses can be generated in the yeast *Saccharomyces cerevisiae*. ***J. Biotechnol.,*** submitted. 2004.

Contact: sasnausk@ibt.lt

Laboratory of Recombinant Proteins Research

Head Dr. Vytautas Naktinis

Gervydas Dienys, Prof., Ph.D.

Gintas Žvirblis, Ph.D.

Henrikas Šebėka, Ph.D.

Asta Zubrienė, Ph.D.

Mindaugas Zaveckas, M.Sc.

Editas Mištinienė, M.Sc.

Milda Plečkaitytė, M.Sc.

Iiga Matijošytė, M.Sc.

Vilma Michailovienė, M.Sc

Birutė Juškaitė, M.Sc.

Kristina Rekstytė, B. Sc.

Neringa Laurikietytė, B.Sc.

Regina Basiul-Špokienė B.Sc

The project was implemented in coordination with the R&D centre of the spin-off company UAB SICOR Biotech involved in recombinant proteins as pharmaceuticals producing and directed to basic research in the areas important to the company. Some enzymatic investigations were also carried-out in collaboration with UAB SICOR Biotech and Vilnius University.

Investigation of quaternary structure and ligand binding of proteins UK114 and p14.5

Both proteins UK114 and p14.5 are members of the putative family of small proteins YER057c/YIL051c/YjgF. The amino acid sequences of the members of this family are highly conserved throughout the evolutionary tree, from prokaryotes to eukaryotes. Despite a conservation and high degree of identity between structures the biological role of these proteins is not well understood, and in addition their oligomeric structure in solution remains controversial. UK114, the protein from goat liver is tumour associated antigen (Bartorelli, et. al., 1996). This protein is interesting for pharmaceutical purpose, because a preliminary evidence on an anti-cancer efficacy of UK114-containing preparations in humans has been reported (Mor et. al., 1997). UAB SICOR Biotech considers recombinant UK114 and its analogs as potential anticancer pharmaceuticals.

Not only amino acids sequences of members of YER057c/YIL051c/YjgF family share high identity, but also 3-D structures are very similar. Up to now crystal structures of two bacterial homologues are available: a purine regulatory protein from *Bacillus subtilis* (Sinha et.al., 1999) and YjgfF from *E. coli* (Volz et. al., 1999). Both proteins are homotrimers. Trimeric structure in solution was confirmed for only one bacterial homologue HI0719 from *Heamofilus influenza* (Parsons et. al. 2003). Our previously works clearly demonstrated that the mammalian protein UK114 also exhibited trimeric structure in solution and in crystal. Despite the multifunctionality of the proteins from YER057c/YIL051c/YjgF family, high similarity in quaternary structures suggest that these proteins use the same biochemical way for the manifestation of their biological function.

Our aim was to test one possible mode of action of these proteins, which is related with binding of small ligands. This presumption was supported by identification of three potential ligand binding cavities between interacting subunits, very conserved in all members of YER057c/YIL051c/YjgF family. We tested ligand binding activity for proteins UK114 and p14.5 with free fatty acid and fluorescence probe 8-anilino-1-naphthalenesulfonic acid. We confirmed that both proteins were able to bind hydrophobic ligand with stoichiometry 1:1 mol of ligand per mol monomeric protein.

Our study was the first attempt to describe the structure based ligand binding sites of protein p14.5 from YER057c/YIL051c/YjgF family. From the crystal structure of UK114, very close homologue of p14.5, we identified most conserved amino acids, responsible for the interaction between subunits in the trimer. These amino acids are highly conserved through the whole proteins from YER057c/YIL051c/YjgF family. Using site specific mutagenesis we changed these amino acids in monomer:monomer sites of interaction. Quaternary structure of p14.5 and its mutants was determined using a number of methods. Subsequent ligand binding activity was tested for native homotrimeric p14.5 and monomeric mutants.

Our findings provided a proof of the importance of the trimeric protein structure for its ability to bind hydrophobic ligands. Amino acids responsible for the ligand binding and specificity were also identified. Further we plan to identify possible ligands for UK114 and p14.5 proteins and determine a possible biological role of YER057c/YjgF/UK114 proteins.

Investigation of IFNα2b aggregate formation

Recombinant alpha interferons (IFNα) and beta interferon (IFNβ) belong to the type I interferons, which are the most effective antivirus agents among cytokines. These proteins lose their biological activity if their molecules aggregate into dimers or higher oligomers. Aggregates are also mainly responsible for immune reactions of IFNα.

No much information is available about the process of aggregation. In the case of IFNα it was investigated mainly by immunological methods. Oligomers were found even in the best known in the word market pharmaceutical formulations of IFNα (Braun, A. et al., 1997). Analysis of the new patents show that some components in the solution increase resistibility of IFNα to aggregation. A composition including benzyl alcohol is proposed for IFNα2a (US patent 5.762.923,1998); for IFNα2b - a composition with metal complexes

forming compounds (US patent 5.766.582,1998). Aggregation of IFNβ is suppressed by cyclodextrins (US patent 5.730.969,1998). Various antimicrobial substances are included into the composition of pharmaceutical formulations. There is no information about their influence on the aggregation process. Investigation of aggregate formation is an urgent problem, important for an oriented selection of pharmaceutical formulations components.

Starting investigation of this problem we made an assumption that aggregation of IFNα2b takes place as a result of partial denaturation under the influence of external factors. The rate of denaturation process may be determined from fluorescence spectra of Trp and Tyr residues. The rate of aggregation by method of high pressure chromatography (HPSEC). All experiments were performed in model conditions at elevated temperature (50°C).

Experiments showed that, initially, denaturation changes of exponentially decreasing intensity took place. These changes are well described by the first order kinetic equations. Denaturation is slowest in neutral solution (pH 6-7). It gets 2-4 times more rapid in acidic or alkaline sides of the pH interval 3-10. Initially, denatured IFNα2b molecules aggregate according to the second order kinetics. Later the rate of aggregation changes to the first order kinetics. Aggregation of IFNα2b in the presence of antimicrobial substances was also studied. Their influence fitted pseudo-first order kinetics. Rate of aggregation was 2-9 times higher than in the absence of antimicrobial substances. Aggregation rate constants were strongly dependent on pH in the case of antimicrobial substances containing ionogenic groups in their molecules. Our results are consistent with the initial assumption that the rate limiting stage of aggregation is unfolding of the protein.

Further it would be desirable to search for substances inhibiting aggregation of IFNα2b and study their influence on its denaturation.

Investigation of the interaction of rhG-CSF with immobilized metal ions and the metal ion assisted protein refolding

Expression of recombinant proteins in *Escherichia coli* often results in the formation of insoluble inclusion bodies (Marston, 1986). The recovery of native soluble protein from inclusion bodies involves inclusion body solubilization and refolding of the solubilized protein. *In vitro* protein refolding is still a bottleneck in both structural biology and in the development of new biopharmaceuticals, especially, for commercially important proteins that are overexpressed in *E. coli*.

Recently, aqueous two-phase systems have been applied for protein refolding (Forciniti, 1994). The aqueous two-phase systems, similarly to the other chromatographic processes, may be an attractive alternative for the existing protein refolding methods, because protein refolding can often be combined with simultaneous partial purification. We have successfully applied aqueous two-phase systems polyethyleneglycol/dextran for the refolding of recombinant human granulocyte colony stimulating factor (rhG-CSF), its Cys17Ser and hexahistidine-tagged forms solubilized from inclusion bodies upon their binding to the immobilized Ni(II) and Cu(II) ions (Zaveckas et. al, 2000).

Immobilized metal ion affinity-based techniques were successfully used for refolding of proteins equipped with engineered polyhistidine tags, however, naturally-occuring metal binding sites have received little attention. Refolding upon binding metal ions via naturally existing metal binding residues may be applicable to the other proteins, however, it is necessary to reveal which metal binding residues dominate in the interaction with immobilized metal ions. The studies of rhG-CSF, rhG-CSF (C17S) and hexahistidine-tagged rhG-CSF interaction with immobilized Cu(II), Ni(II) and Hg(II) ions revealed for the first time the participation of Cys17 residue in the protein interaction with chelated Hg(II) ions, and the prevalence of histidine residues in the protein interaction with Cu(II) and Ni(II).

Our study intended to find the effect of histidine mutations of rhG-CSF on protein interaction with metal ions and metal ion-assisted refolding. The refolding and interaction of rhG-CSF His^{43} and His^{52} mutants with immobilized Ni(II) and Hg(II) ions in aqueous two-phase systems was investigated. This study revealed the crucial importance of histidine His^{52} residue on protein interaction with Ni(II) ions and Ni(II) ion-assisted refolding. Possibly, Ni(II) ion binds to the neighbouring His^{52} and His^{156} residues.

Further studies on rhG-CSF histidine mutants interaction and refolding in the presence of chelated metal ions could help us to reach more strict conclusions regarding the rhG-CSF interaction with metal ions and metal ion assisted refolding, and the application of this refolding method for other proteins containing naturally occurring metal binding residues.

Grants

Lithuanian State Science and Studies Foundation

Contracts

UAB Sicor Biotech, Lithuania, Vilnius

Collaboration

Vilnius Gediminas Technical University, Lithuania

Recent publications

Zubrienė A., Budrienė S., Lubienė J. and Dienys G. (2002). Immobilized alkaline phoshatase for molecular cloning. *Biocatalysis Biotransform*. 20(6):423-427.

Rožėnaitė V., Baškevičiūtė B., Lukša V., Bumelis V., Pesliakas H. (2002). Immobilized metal ion affinity refolding of recombinant proteins. ***Biologija*** 4:25-29.

Chmeliauskaitė V. and Žvirblis G. (2002). Cloning and expression of *Xenopus laevis* chaperones Hsp30C and Hsp70 in *E. coli*. ***Biologija*** 4:42-44.

Mistiniene E., Lukša V., Sereikaite J. and Naktinis V. (2003). Oligomeric assembly and ligand binding of the members of protein family YER057c/YIL051c/YjgF. ***Bioconjugate Chem***. 14(6):1243-1252.

Deriu D, Briand C., Mistiniene E., Naktinis V. and Gruetter M.G. (2003). Crystal structure and oligomeric state of the mammalian tumor associated antigen UK114. ***Acta Crystallogr D Biol Crystallogr*** 59(Pt 9):1676-1678.

Zaveckas M., Lukša V., Žvirblis G., Chmeliauskaitė V., Bumelis V. and Pesliakas H. (2003). Mutation of surface-exposed histidine residues of recombinant human granulocyte-colony stimulating factor (Cys17Ser) impacts on interaction with chelated metal ions and refolding in aqueous two-phase systems. ***J. Chromatogr. B*** 786:17-32.

Dienys G., Jarmalavičius S., Budrienė S., Čitavičius D., Sereikaitė J. (2003). Alcohol oxidase from the yeast *Pichia pastoris* - a potential catalyst for organic synthesis. ***J. Mol. Catalysis B: Enzymatic*** 21:47-49.

Zubrienė A., Budrienė S., Gorochovceva N., Romaškevič T., Matulionis E. and Dienys G. (2003). Immobilization of hydrolases onto chitosan microparticles. ***Chemija*** 4:226-230.

Pleckaityte M., Mistiniene E., Michailoviene V. and Žvirblis G. (2003). Identification and characterization of Hsp70 (DnaK) chaperone system from *Meiothermus ruber*. ***Mol. Gen. Genomics*** 269:109-115.

Contact: gervydas.dienys@chf.vu.lt

Laboratory of Immunology

Head Dr. Aurelija Žvirblienė

Arvydas Kanopka, Ph.D.

Violeta Popendikytė, Ph.D.

Petras Stakėnas, Ph.D.

Eglė Markauskaitė, M.Sc.

Daiva Bakonytė, M.Sc.

Aurelija Baranauskaite, M.Sc.

Loreta Jasiulevičiūtė, M.Sc.

Gražina Petrauskienė, M.Sc.

Natalija Pozdniakovaitė, M.Sc.

Laura Šamonskytė, B.Sc.

Leokadija Diglienė

The Laboratory of Immunology involves several research groups focused to cytokine research, hybridoma technology (Dr. A.Žvirbliene), epigenetic studies of genes involved in complex diseases (Dr. V.Popendikytė), regulation of gene expression by alternative splicing (Dr. A.Kanopka), molecular epidemiology of tuberculosis (Dr. P.Stakėnas). The Laboratory has basic infrastructure for tissue culture, molecular biology, immunochemical and immunofluorescence analysis. It maintains good contacts with hospitals and clinical laboratories in Lithuania.

Studies on the role of cytokines in chronic inflammation. Development of new monoclonal antibodies

Since 1996, the Laboratory of Immunology is involved into collaborative studies aimed to immunopathology mechanisms in oncologic and infectious diseases. In 2001-2003 our research group together with Vilnius University Central Hospital participated in a collaborative project on immune response mechanisms in chronic inflammation. The Lithuanian Science and Studies Foundation supported this study.

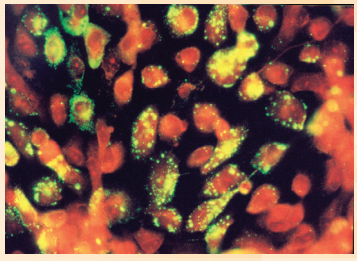
The aim of this study was to analyse cytokine production in patients with acute and chronic reactive arthritis (AcReA /CReA) in order to better understand the pathophysiological mechanisms of CReA and search for new treatment possibilities. Reactive arthritis (ReA) is an acute inflammatory disease, which develops into severe, chronic form in approximately 15-30% of patients. Pathogenesis of ReA is closely associated with an earlier bacterial infection. Molecular mechanisms responsible for the chronic and destructive course of ReA are not fully understood. Cytokines secreted by T helper (Th1/Th2) cells are indispensable for the effective defense against bacterial infection. In ReA, impaired Th1 cytokine production may cause a failure of effective bacterial elimination at the initiation of the disease and thereby participate in pathogenesis of CReA. However, data on the role of Th1/Th2 cytokines in the chronic outcome of the disease are controversial. This lack of general agreement on cytokines found in ReA patients warrants further cytokine investigation in ReA.

The objective of this study was determination of cytokine-producing T cell populations and *in vitro* secreted cytokine profiles of patients with ReA in comparison with rheumatoid arthritis (RA) patients and healthy donors. For that purpose, production of tumour necrosis factor alpha (TNF-α), interferon-gamma (IFN-γ) and interleukin-10 (IL-10) was measured by ELISA in peripheral blood mononuclear cell (PBMCs) cultures from 28 patients with AcReA, 27 patients with CReA, 26 patients with RA and 31 healthy control persons. The percentages of TNF-α, IFN-γ, and IL-4 positive CD3+ T-cells were determined in the same groups of patients and healthy donors by means of flow-cytometry. In addition, cytokine production by synovial fluid cells of patients with AcReA, CReA and RA were evaluated.

Comparison between these groups of patients and also healthy control group allowed a comprehensive evaluation of changes in cytokine profiles related to chronicity of ReA. Our study revealed significant differences in cytokine profiles in CReA and AcReA. This was indicated by a higher production of TNF-α observed in PBMCs of patients with CReA than with AcReA. In addition, the percentage of IFN-γ+ blood and synovial fluid T cells was significantly higher in CReA than in AcReA. It is interesting to note that no difference in Th2 cytokine production (IL-4 and IL-10) was observed when CReA was compared to AcReA.

Our data support the idea that the relative lack of Th1 cytokines may play a significant role in pathogenesis of ReA at the acute phase of disease. However, in the chronic phase of ReA the cytokine repertoire is different and similar to that found in RA. Enhanced TNF-α expression observed in the chronic phase of disease indicates the importance of this cytokine in the pathophysiological events of CReA and may suggest TNF-α as a potential therapeutic target in CReA.

In 2002-2003, our research group was also involved in the production of new monoclonal antibodies using hybridoma technology. In collaboration with the Laboratory of Eukaryote Genetic Engineering that provided recombinant yeast-expressed proteins, a large panel of new hybridoma cell lines producing monoclonal antibodies to viral antigens were generated. Monoclonal antibodies are useful diagnostic tools. They were successfully used for the development of highly sensitive and specific immunoassays to detect mumps and measles infection.



Immunofluorescence staining of mumps virus infected Vero cells using monoclonal antibody specific to mumps virus nucleocapsid protein.

Epigenetic Analysis of Class iii ins vntr Alleles

During last six years several research projects were performed in the Laboratory of Immunology collaborating with The Krembil Family Epigenetics Laboratory, Centre for Addiction and Mental Health (University of Toronto, Canada). Research projects were dedicated to understanding of DNA methylation and its potential role in gene control during development of complex non-Mendelian diseases such as schizophrenia, miotonic dystrophy, Huntington disease, diabetes.

Human morbid genetics, a field that just ten years ago nearly exclusively dealt with rare classical Mendelian genetic disorders, now invests its major resources into uncovering the molecular genetic basis of wide scope of complex diseases. In the overwhelming majority of cases evidence for association to a disease derives from haplotypes that are built on single nucleotide polymorphisms (SNPs) of unknown functional significance, most often such SNPs do not change amino acid sequence and/or affect expression or splicing of the gene. As a rule the associations of the identified risk haplotypes are not of sufficient magnitude to fully explain the linkage data. Traditional linkage and association studies have turned out to be very controversial, and new ways of analysis are required.

One of the recent developments in psychiatric and other numerous complex diseases research is psychiatric epigenetics. By definition, epigenetics refers to modifications in gene expression that are controlled by heritable, but potentially reversible changes in DNA methylation and/or chromatin structure. In the last several decades, a myriad of epigenetic mechanisms has been uncovered, and a series of important functions of genome regulation have been ascribed to epigenetic DNA modification. A large proportion of genes exhibit correlation between gene expression and the degree of methylation of cytosine residues, i.e. the lower the degree of DNA methylation in the gene regulatory region, the higher the expression of the gene is expected. Although the cause-effect relationship between epigenetic status of a gene and gene expression has been questioned, there are experimental data supporting the idea that epigenetic factors play a gene regulatory role. DNA methylation and chromatin conformation are two anatomically different but functionally related molecular mechanisms of epigenetic regulation.

To analyse the biological role of 5-methylation of cytosine residues (5-MeCs) in DNA requires precise and efficient methods for detecting individual 5-MeCs in genomic DNA. Last ten years modern epigenetic studies applied the technique of direct sequencing of sodium bisulphite treated DNA for mapping of methylated cytosines.

Type 1, or insulin-dependent, diabetes mellitus is a complex non-Mendelian disease, the key pathogenic mechanism of which consists of autoimmune T lymphocyte- mediated destruction of insulin-producing cells. Several case-control studies detected evidence for association between polymorphisms of the gene encoding insulin (*INS*) and type 1 diabetes. The results of case-controls studies were replicated in family-based genetic studies. A cross-match haplotype analysis of 10 polymorphisms at *INS* locus suggested that the type 1 diabetes susceptibility site corresponds to the variable number tandem repeat (VNTR) region at the *INS* promoter region and this locus was named *IDDM*2. The *INS* VNTR represents a polymorphic minisatellite embedded in the promoter region, 600 bp upstream from the *INS* transcription start site. The VNTR is composed of tandemly repeated sequences of 14-15bp oligonucleotide. The repeat number varies significantly, and the VNTR alleles are grouped into two dominating classes: class I (26-63 repeats), and class III (141-209 repeats), with frequencies in Caucasians of 0.71 and 0.29, respectively. It has been consistently detected that VNTR class I alleles are associated with type 1 diabetes, and the class I/I homozygous genotype increases risk for type 1 diabetes 2-5- fold. Class III alleles are protective, and a single copy of such allele reduces the risk to the disease by 60%-70%.

Thymus represents a critical site for tolerance induction to self-proteins. Immature T cells with potential autoreactivity to self-proteins, such as insulin, are normally destroyed in the thymus. Reduced expression of *INS* in thymus during embryogenesis may lead to insufficient deletion of the insulin-specific T-cells which eventually may attack pancreatic islets. Two studies detected that class III VNTR alleles were associated with 2- to 3-fold higher *INS* mRNA levels than the class I alleles in the human embryonic and post-natal thymi. It is interesting to note that five class III alleles out of 25 class I/III heterozygous samples exhibited no expression in thymus at all, which implies that such alleles are predisposing to type 1 diabetes in a similar manner to class I alleles. These alleles were named 'S' (silencing) alleles. Sequencing of an "S" allele and one of the abundantly expressing class III allele (called 'enhanced insulin expression allele', or "E") detected only small sequence differences between the "E" and the "S" alleles.

Epigenetic studies are highly relevant to the further understanding of the mechanism how *INS* VNTR alleles predispose or protect from type 1 diabetes. Parent of origin effects at the *INS* VNTR may be related to genomic imprinting, an epigenetic phenomenon where

by genes are expressed or not expressed depending on their parental origin. *INS* is positioned in the cluster of imprinted genes on chromosome 11p15.5, and surrounded by imprinted genes *ASCL2, IGF2*, and *H19*. Based on the common observation that promoter hypermethylation leads to gene inactivation, it is very likely that the differences in the mRNR levels of insulin gene VNTR class I and class III alleles are predetermined by different degree of methylation.

The main task of the current research - epigenetic analysis of class I and III *INS* VNTR alleles was performed on DNA samples from human thymus embryos, with VNTR I/III heterozygous genotype (DNA samples were kindly provided by Dr. C. Polychronakos from McGill University-Montreal Children's Hospital Research Institute and Dr A. Pugliese from Diabetes Research Institute, University of Miami, Florida, USA). DNA epigenetic analysis protocols consisted of the following steps: bisulphite modification of DNA samples by treating with 3M Na metabisulphite solution at 55°C for 4h; amplification the region of interest and cloning PCR products into a vector pUC57/T; plasmid DNA extractions and sequencing of inserts. Usually, 10-15 clones from each DNA sample were subjected to sequencing because of the so-called epigenetic mosaicism, i.e. DNA from the cells of the same tissue exhibit epigenetic differences.

Seven human thymus DNA samples, pancreas and leucocytes DNA samples have been used for detection of 5-MeCs positions in 5'-untranslated region and first exon of *INS* gene (distance of 3420 nt.). Data about insulin gene VNTR alleles methylation were collected from sequences of 435 clones. Sequences of different *INS* gene VNTR alleles were distinguished according to linkage disequilibrium of polymorphic HphI site near *INS* gene translation start (at -23 nt.).

DNA methylation differences have been detected in the sequences of insulin gene VNTR class I and class III alleles (two thymus DNA samples) 5-MeCs have not been found in the sequences of insulin gene VNTR class III alleles in the distance between 4309 nt. and 4613 nt., (distance of 300 nt.), with the exception of 5-MeCs at 4491 nt. Each cytosine at CpG position was methylated in the sequences of insulin gene VNTR class I alleles. The distance of sequences between 4309 nt and 4485 nt involves promoter sequence and part of the first exon of *INS* gene. Preliminary data of epigenetic analysis confirmed the main hypothesis of the scientific research: alleles of insulin gene-VNTR class III and class I are differently methylated in human thymus. Methylation differences have been detected only in two DNA samples from seven used for that pilot research study. Possibly, these DNA methylation differences may explain in some degree the differences in the insulin gene expression depending on the length of VNTR near insulin gene, but real estimation of epigenetic regulation in type 1 diabetes needs additional extensive scientific research.

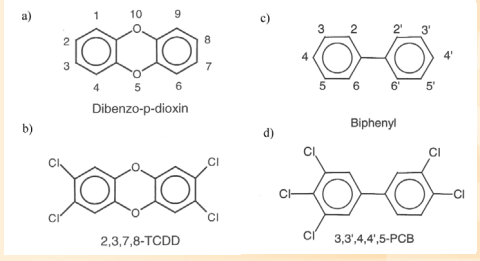
Comprehensive risk analysis of dioxins: development of methodology to assess genetic susceptibility to developmental disturbances and cancer

An *in vitro* splicing system as an aid in extrapolation of animal data to humans

Dioxin risk assessment is one of the most controversial issues in the history of risk assessment. There are many reasons for this. Some of these compounds are highly toxic or „supertoxic" in animal experiments. Public interest and fear arose from several incidents.

In 1976 a trichlorophenol plant in Seveso, Italy, released a remarkable amount of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) to the environment, and the town was evacuated, and the area and the population have been monitored ever since. In 1968 and 1979 Yusho and Yu-Cheng accidents took place in Japan and Taiwan, respectively. The toxic ingredient is assumed to be mainly polychlorinated dibenzofuran impurities in polychlorinated biphenyl (PCB) oil, which leaked to rice oil, and it caused a number of toxic effects including foetal death and developmental effects. In 1999 chicken and pork feed were found to be contaminated by dioxins and PCBs, and this caused huge economic losses due to withdrawal from the market of the whole productions of farms using animal feed from several feed producers.

Dioxins bring about a wide spectrum of biochemical and toxic effects in experimental animals. These effects depend on species, strain, gender, age and tissue. In humans, a wide variety of health effects have been linked to exposure to dioxins, including mood alterations, reduced cognitive performance, diabetes, changes in lymphocyte subsets, dental defects, endometriosis, decreased male/female ratio of births and decreased testosterone and (in neonates) elevated thyroxin levels.



Structures of dibenzo-p-dioxin (a), TCDD (b), biphenyl (c) and PCB (d)

A common denominator of dioxin effects appears to be the AH receptor (AHR), which mediates the induction of CYP1A1 enzyme and probably also most other biological effects of TCDD.

Acute toxicity is very different in different animal species, e.g. guinea pig is over thousand-fold more sensitive than hamster. At doses that are also otherwise toxic, they have been shown carcinogenic in several animal species. They are also clearly fetotoxic and teratogenic in animal experiments, and recently more subtle developmental effects have raised concern. Recently also data have been published showing variation of toxicity within a single species. It has been demonstrated a striking difference in susceptibility to the acute lethality of TCDD between two rat strains, inbred Long-Evans (L-E) and outbred Han/Wistar (H/W). The LD50 values for these strains are about 10 and >10,000 ng/kg, respectively. Thus, the H/W rat is the most TCDD-resistant mammal known while L-E rats are the most TCDD-sensitive rats ever tested. The differences between L-E rats and H/W rats in sensitivity to TCDD correlate with differential AHR mRNA processing between L-E rats and H/W rats at the exon 10 region and that these differences are due to an invariant GT nucleotide at the end of intron 10.

The aim of the work was to establish an *in vitro* mRNA splicing system to enable us to compare splicing profiles from L-E and H/W rat AHR pre-mRNAs to those of human AHR pre mRNA populations, in particular, containing mutations in the intron 10/exon border region. To evaluate consequences of aryl hydrocarbon receptor (AHR) mutations, a pre-mRNA splicing model constructs of aryl hydrocarbon receptor (AHR) from rat wild type (L-E) and mutant (H/W) AHRs were generated. These constructs were tested in an *in vitro* splicing reconstitution reaction in HeLa cell nuclear extracts, and products of the splicing reaction were analysed on 8% denaturing polyacrylamide gels. According to the result, the use of the AHR construct from H/W rat (mutated 5' splice site) revealed the usage of cryptic splice site, which normaly is not used for the whild type rat AHR. Thus, using *in vitro* system performed in human cells we have reconstituted pre-mRNA splicing events which occurred in rats.

The comparison of the splicing patterns of AHR pre-mRNAs may allow us to discover consequences of potential human AHR mRNA variants in which sequences encoding the C-terminal transactivation domain of the AHR are processed. For this purpose, control and tumor-derived human tissue samples from liver, lung, stomach, kidney, mammilla, mucous and pancreas were collected (kindly provided by the Lithuanian Centre of Oncology). DNA was isolated from the samples and a 10-intron region of the human AHR (hAHR) gene was amplified by polymerase chain reaction (PCR). PCR products were sequenced and normal and tumor DNA sequences, encompassing the 10-th intron region of AHR gene, were compared. The results revealed that, in polypyrimidine tract sequence, which is directly adjanced to 3' splice site, there were transitions of T to A observed in few cases in lung cancer tissues. Since the polypyrimidine tract is „strong", in splicing terms, it is hard to believe that a single mutation can affect binding of splicing factors to pre-mRNA, and, thus, influence the processing of AHR gene. Comparison of other tested cancer tissues no changes were observed in the studied pre-RNA splicing region of the AHR.

Our results demonstrate that, at genomic level, probably there are no differences between normal and cancer tissues observed in the intron 10 region of AHR, that might effect AHR pre-mRNA processing.

Of course, we did not test all cancer types (phenotypes), thus we could not exclude the possibility that there were no changes in AHR intron 10 region in all human cancer types at all. To answer this question, the mutation analysis of AHR mRNA in different types of human cancer should be done in further studies and these experiments are critical in order to refine the model of human risk assessment.

Molecular epidemiology of *Mycobacterium tuberculosis* in Lithuania

Tuberculosis (TB) continues to be a serious global health problem. An estimated one third of the world's population is infected with *M. tuberculosis* complex bacteria causing 9 million of new cases and 3 million deaths annually. Most effectively tuberculosis is cured according to WHO recommended directly observed treatment short-course (DOTS) strategy that includes treatment regimes using three or four essential anti-TB drugs. However, an improper management of tuberculosis has led to emergence of multi-drug-resistance (MDR), defined as resistance to the two most powerful anti-tubercular drugs isoniazid (INH) and rifampin (RIF) at least. According to DOTS-plus strategy, MDR-TB is treated by using the second-line anti-tubercular drugs and in the developed countries between 60 and 80% of patients with MDR are cured. However, such treatment is far more expensive and in developing nations most patients with

MDR-TB is condemned to die. Moreover, the average MDR-TB patient may infect a further 15 to 20 people. Therefore, it is of great importance to improve the understanding of the transmission of tuberculosis and the mechanisms of acquisition of drug resistance of *M. tuberculosis*.

Until the year 1989, tuberculosis in Lithuania was kept under control and number of cases was dropped to 39.5 per 100 000 inhabitants. However, during the last years due to a socioeconomic difficulties and inadequacies of tuberculosis control policy, incidence of tuberculosis increased in twice, reaching a peak of the total notified 3176 cases in 1998. There were 2097 notified new and 323 recurrent TB cases in 2002, notification rate was 74.7 per 100 000 of population. The emergence of drug-resistant tuberculosis is one of the highest in the world. In 2002, 9.4% of *M. tuberculosis* isolates recovered from the new and 53.3% isolates from treated previously TB cases were MDr. Without any doubts, there are strong reasons for increased efforts, including scientific research efforts, to counteract the threatening situation. In this context, it is urgently needed to improve diagnostic capabilities for the detection and identification of infection agent, to define spectrum of mutations in *M. tuberculosis* responsible for the drug-resistance and to get an objective overview on the existence, arising and clonal spread of the most virulent strains circulating in Lithuania.

To meet these requirements the Institute of Biotechnology started integration of molecular approaches into conventional studies of tuberculosis under ongoing project *Molecular Epidemiology of Tuberculosis in Lithuania* in the end of 1998. The research has been carried out in collaboration with the Republican hospital of tuberculosis and infectious diseases.

In summary, *M. tuberculosis* complex IS*6110*-specific Polymerase Chain Reaction (PCR) assay was evaluated in a clinical study by investigating 200 patients with suspected tuberculosis. The results of this study confirmed that PCR technique is especially useful diagnostic tool for the detection of culture positive tuberculosis and is promising one when culture's test failures.

Spectrum of mutations associated with *M. tuberculosis* resistance to isoniazid was investigated by direct sequencing and the developed PCR-based assay. Mutations at codon 315 of the *katG* gene were detected in 86% of 364 INH-resistant *M. tuberculosis* isolates. Fourteen percents 52 INH-resistant isolates with wild type codon Ser315 and 19% of 52 INH-resistant isolates with mutated *kat*G allele had a single mutation in the promoter of *mabA-inhA* operon. It was demonstrated for the first time that frequency of Ser315 mutations depends significantly on a drug resistance pattern among IHN-resistant *M. tuberculosis* isolates. The results of this study justified a simple PCR-based assay for prediction of INH-resistant *M. tuberculosis* in Lithuania and indicated that majority of INH-resistant cases had undergone multiple rounds of inadequate therapy.

In total, 196 rifampin-resistant *M. tuberculosis* isolates were searched by sequencing for the mutations in the *rpoB* gene in aim to evaluate significance of these mutations for the rapid prediction of drug-resistance in Lithuania. The mutations in hot spot cluster I region were identified in 94% isolates. Seventeen different types of mutations involving nine codons were identified. No any of silent mutation was found. The majority of them (96%) were single point mutations. Most frequently mutations were observed in codons Ser531 (46%), Asp516 (21%), and His526 (19%). These results indicated that mutations observed among RIF-resistant *M. tuberculosis* isolates are consistent with distribution tendencies of mutations observed among isolates from the majority of other countries. Therefore, the implementation of yet available genotypic tests would be especially useful for the rapid identification of *M. tuberculosis* resistance to rifampin and prediction of multi-drug-resistance in Lithuania.

The PCR-based technique for the rapid detection of mutations in the codons 43 and 88 of rpsL gene leading to acquisition of *M. tuberculosis* resistance to streptomycin (STR) was developed and introduced to investigate prevalence of these mutations among *M. tuberculosis* strains circulating in Lithuania. Out of 468 STR-resistant isolates tested, the mutations in codons 43 and 88 of *rpsL* gene were detected in 46% and 12% of isolates, subsequently. These results demonstrated that developed assay could serve as promising diagnostic tool for the rapid prediction of *M. tuberculosis* resistance to this first-line drug in half of STR-resistant cases. Search for mutations in the *rrs* gene that also are associated with *M. tuberculosis* resistance to streptomycin is in progress.

In total, 642 *M. tuberculosis* isolates, including 367 MDR, collected from different patients (338 new, 287 treated previously and 17 with unknown case status) were analyzed by internationally standardized restriction fragment length polymorphism (RFLP) typing using the insertion element IS6110 as a probe. Drug-resistant isolates also were searched for the mutations as described above. Four hundred isolates (62.3%) were found in 74 clusters and 242 RFLP patterns were observed only once each. The size of clusters varied

between two and 44 isolates. Fifty-three percents isolates from new TB cases were in the clusters, while a proportion of clustered isolates from treated previously TB cases was significantly higher (73%, $P<0.001$). Drug-resistant isolates were found in the clusters more frequently than drug-susceptible ones (71% and 38%, $P<0.001$) as well as MDR isolates versus drug resistant but yet not MDR (75% versus 57%, $P<0.001$). Pattern analysis revealed that 66% of isolates, including 78% of all drug-resistant isolates and 86% of all MDR isolates could be attributed to four large families of related strains. Remarkably, these families had the specific patterns of mutations associated with drug resistance in the genes investigated. The largest family (161 isolates, 44 strains) according to the profiles of IS*6110* RFLP patterns could be allocated to the Beijing genotype, which was confirmed by sub-typing of these isolates using alternative PCR-based approach.

The results of this study revealed threatening TB situation in Lithuania. In general, clustered isolates represent a recent transmission. Fifty-three percent of searched isolates from new TB cases were in the clusters. The rate of clustering among isolates from new cases was extremely high considering that number of searched isolates comprised approximately only three percent of the new cases notified during this study. Furthermore, a frequency of clustered isolates among isolates from treated previously patients was significantly higher than that among isolates from new TB cases. The rate of clustered drug-resistant isolates in comparison with drug-susceptible isolates was twice higher as well. In addition, the clustering also was less among drug-resistant isolates but yet susceptible to RIF or INH, i.e., among the cases with a better outcome of the treatment than among MDR cases. Moreover, many of isolates from new cases that fell into clusters shared with isolates from previously treated patients were MDr. In sum, these data indicated on serious misrule in management of tuberculosis leading to a frequent infection from uncured patients. Furthermore, the discovery of four large families of genetically related strains circulating in Lithuania is the earliest dangerous sign that also other than Beijing genotype highly transmissible families of drug-resistant *M. tuberculosis* could spread in Europe.

The search for re-infection of *M. tuberculosis* started with investigation of 16 patients suffering from chronic or recurrent disease. No any case of re-infection was found, thus failure of the initial treatment was demonstrated. The majority of these patients were infected by *M. tuberculosis* belonging to the strains of genetically related groups mentioned above. These results once again indicate the necessity to improve quality of treatment and epidemiological control of tuberculosis in order to interrupt transmission chains of the most virulent strains.

Future prospects include: a) monitoring of *M. tuberculosis* strains circulating in Lithuania, b) introduction of data on the Lithuanian genotypes of tubercle bacillus into international databases and search for international transmission of the most virulent strains/genotypes, c) implementation of molecular epidemiology approach into the treatment of MDR tuberculosis. The finalizing goal would be to establish National drug-resistant tuberculosis network to monitor the spread of tuberculosis and quickly alert national and EU institutions about the occurrence of common source outbreaks.

Grants

EU Framework 5 Programme
University of Toronto
Bundesministerium fur Bildung und Forschung
NorFa network
EC INCO-Copernicus Programme
Swedish Institute for Infectious Disease Control
Lithuanian State Science and Studies Foundation

Contracts

UAB Sicor Biotech, Lithuania, Vilnius

Collaboration

Prof. J.Tuomisto Jouko Tuomisto, Department of Environmental Health, National Public Health Institute, Kuopio, Finland

Prof. L.Poellinger, Department of Cell and Molecular Biology, The Medical Nobel Institute, Karolinska Institut, Stockholm, Sweden

Prof. S.Alaluusua, Institute of Dentistry, University of Helsinki, Finland

Prof. P.Mocarelli, Azienda "Ospedale di Vimercate", Ospedale di Desio, University Departament of Clinical Pathology, Desio-Milan, Italy

Prof. P.Soderkvist, University of Linkoping, Department of Biomedicine and Surgery, Linkoping, Sweden

Dr. Sven Hoffner, Swedish Institute for Infectious Disease Control, Solna, Sweden

Prof. Viesturs Baumanis, Biomedical Research and Study Centre, University of Latvia, Latvia

M.D. Anaida Sosnovskaja, Republican Hospital of Tuberculosis and Infectious Diseases, Vilnius, Lithuania

Publications

Petronis A., Popendikyte V, Kan P., Sasaki T. (2002). Major Psychosis and Chromosome 22: Genetics Meets Epigenetics. ***CNS spectrums*** 7:209-213.

Žvingila D., Popendikytė V. (2002.) The genetic diversity of barley cultivars from the Baltic States and Belorussia, as determined by RAMPS. ***Cell Mol. Biol. Lett.*** 7: 477-481.

Makino Y., Kanopka A., Wilson W. J., Tanaka H., Poelinger L. (2002). Inhibitory PAS domain protein (IPAS) is a hypoxia-inducible splicing variant of the hypoxia inducible factor-3α locus. ***J. Biol. Chem.*** 277:32405 – 32408.

Characiejus D., Pašukonienė V., Kazlauskaitė N., Valuckas K. P., Petraitis T., Mauricas M., Den Otter W. (2002). Predictive value of CD8highCD57+ lymphocyte subset in interferon therapy of patients with renal cell carcinoma. ***Anticancer Res.*** 22 (6B):3679–684.

Zvirbliene A., Ranceva E., Jasiuleviciute L., Rudaitiene S., Lazutka J., Mauricas M., Venalis A., Butrimiene I. (2003). Enhanced TNF-alpha expression in chronic reactive arthritis. ***Immunol. Lett.*** 87: 61.

Bakonyte D., Baranauskaite A., Cicenaite J., Sosnovskaja A., Stakenas P. (2003). Molecular characterization of isoniazid-resistant *Mycobacterium tuberculosis* clinical isolates in Lithuania. ***Antimicrob. Agents Chemother.*** 47 (6):2009-2011.

Petronis A, Gottesman II, Kan PX, Kennedy JL, Basile VS, Paterson AD, Popendikyte V. (2003) Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance. ***Schizophr. Bull.*** 29:169-178.

Kan PX, Popendikyte V, Kaminsky ZA, Yolken RH, Petronis A. 2004 Epigentic Studies of Genomic Retroelements in Major Psychosis. ***Schizophr. Bull.***, 67: 95-106.

Butrimiene I., Ranceva J., Jarmalaite S., Venalis A., Jasiuleviciute L. and Zvirbliene A. Different cytokine profiles in patients with chronic and acute reactive arthritis. ***Ann. Rheum. Dis.*** 2004, submitted.

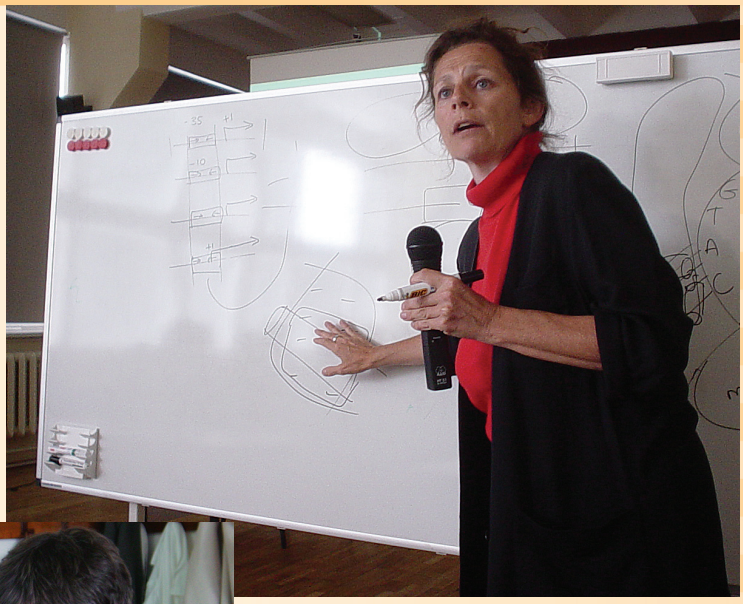
Contact zvirbliene@ibt.lt, stakenas@ibt.lt, kanopka@ibt.lt, popendikyte@ibt.lt

An Intensive Lecture/Laboratory Course on Molecular Interactions of Proteins and Nucleic Acids

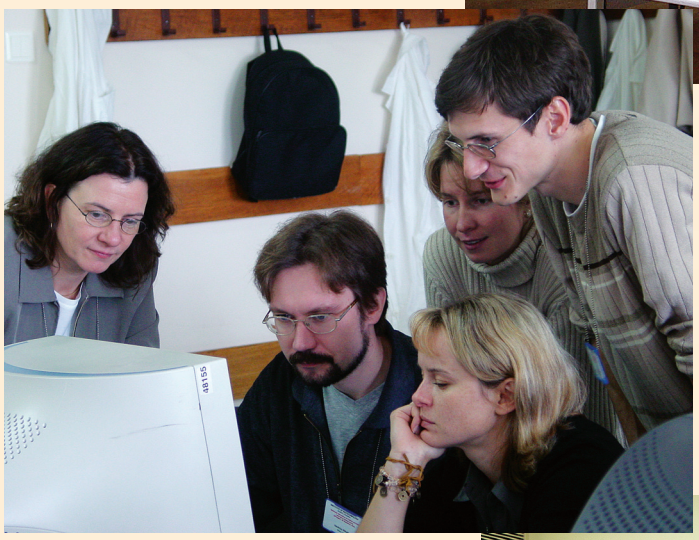
Altogether: students and teachers at the August 25-September 6 course on *Molecular Interactions of Proteins and Nucleic Acids.*

Afternoon with Prof. Geffrey Gardner (the second on the left) from the University of Illinois at Urbana at the seminar on *Integration and Excision of the Conjugative Transposon CTnDOT.*

Prof. Stephen Halford (left) from the University of Bristol and Prof. Bernard Connolly (the second on the left) from the University of New Castle after his seminar on *Fluorescence Methods for Studying Protein-DNA interactions* together with Dr. Virginijus Šikšnys and Mrs. Dalia Daujotytė



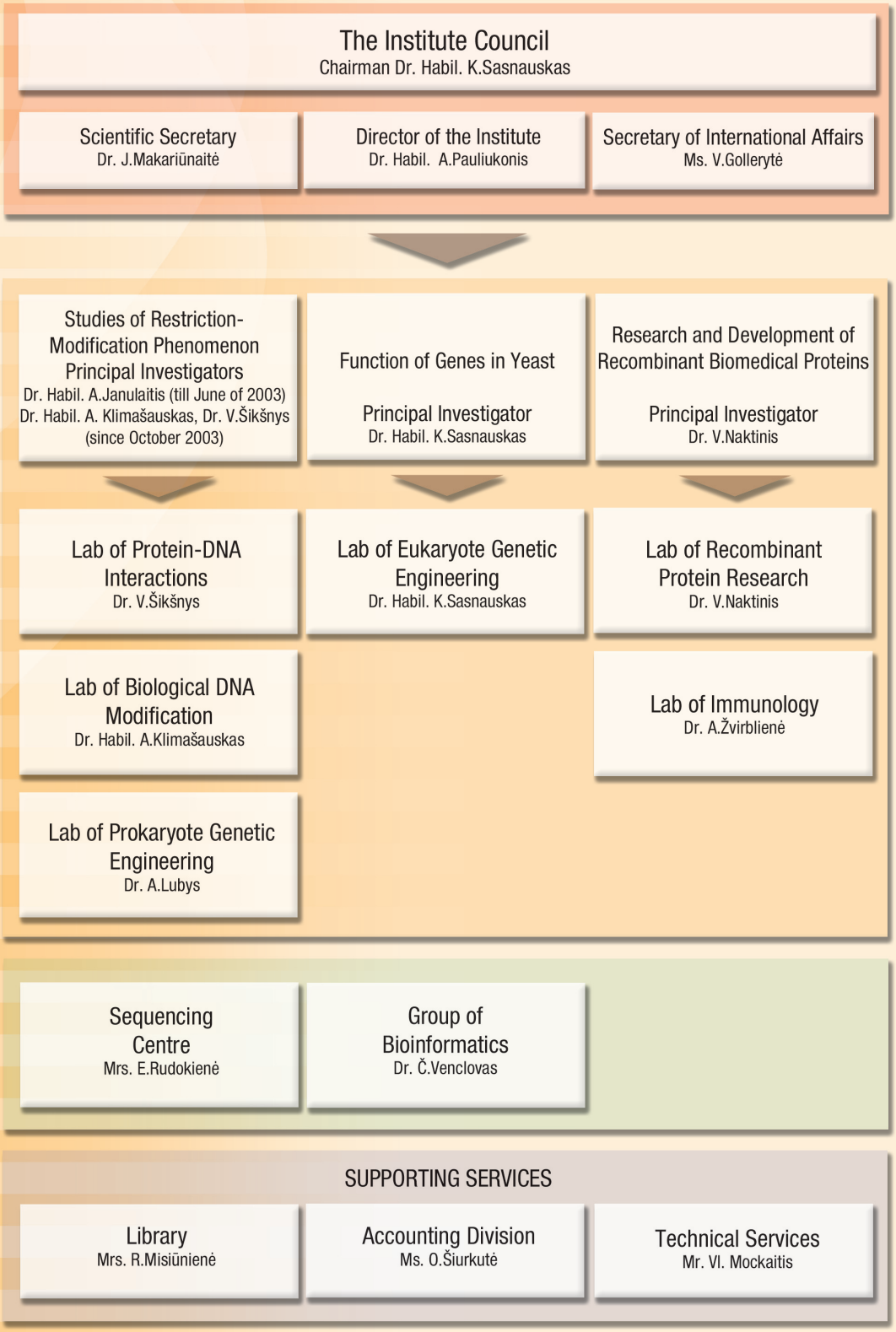
Prof. Jannette Carey from Princeton University presents *The Gestalt of Molecular Recognition.*



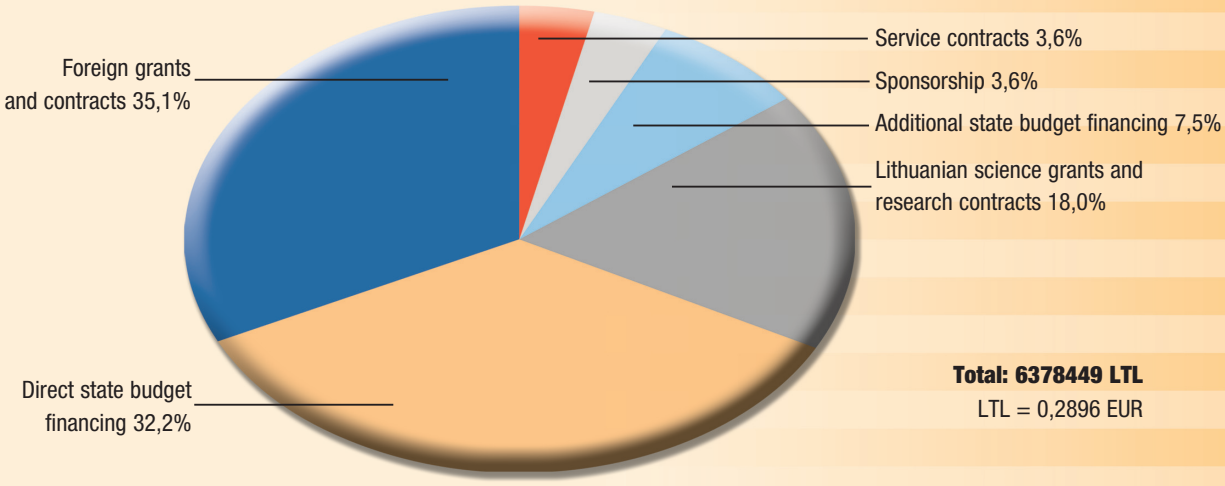
Students actively involved in Prof. Mair Churchill's (from the University of Colorado Health Centre) seminar on *Throwing A Curve at DNA: Structure and Function of Chromosomal HMG Proteins.*

Coordinator of the courses Mrs. Kelley Robbins (left) from the National Academies Washington DC speaks to prof. Richard Gumport (centre) from the University of Illinois at Urbana and Dr. Habil. Saulius Klimašauskas (right) from the Institute of Biotechnology, who headed the courses.

Structure and Management of the Institute of Biotechnology



Financing sources, 2003



Total: 6378449 LTL
LTL = 0,2896 EUR

Foreign grants

- EC FW5 QoL Research project “Bivalent hantavirus vaccine for Europe: Different approaches and evaluation in animal models”.
- EC FW5 QoL Research project “Comprehensive risk analysis of dioxins: development of methodology to assess genetic susceptibility to developmental disturbances and cancer”.
- EC FW5 QoL Research project “Combined immune and gene therapy for chronic hepatitis”.
- EC FW5 QoL Concerted action “Enhanced laboratory surveillance of measles”.
- EC FW5 QoL Research project “Development of highly special enzymes for genome manipulation”
- NATO Science for Peace Programme grant “Natural resources for industry: Investigation of protein refolding factors and their implementation into the biotechnical process”.
- Volkswagen Stiftung, via Max-Planck-Institut für molekulare Physiologie, Research collaboration grant “Rational design and molecular evolution of DNA methyltransferases for new sequence-specific chemical modifications of DNA”.
- The Wellcome Trust Collaborative research initiative grant “Restriction enzymes with novel reaction mechanisms“.
- Howard Hughes Medical Institute International research scholarship grant “Sequence recognition and base flipping by DNA methyltransferases: Structural studies and redesign for novel functions”.
- Howard Hughes Medical Institute Mini Grant to Foster Collaboration Initiative "Towards engineering of restriction enzymes"
- Howard Hughes Medical Institute Mini-Grants To Foster Collaborations “Bioinformatics-guided engineering of DNA methyltransferases”
- Howard Hughes Medical Institute International research scholarship grant “Principles of restriction enzymes specificity”.
- Howard Hughes Medical Institute International research scholarship grant “Combination of improved methods with expert knowledge to derive models of protein structure at low sequence homology”.
- European Science Exchange Programme Joint project grant “Correlating structure and spectroscopy: 2-aminopurine fluorescence from protein-DNA crystals”.
- Nordic Country Cancer Society supported projects on "Nordic network on cellular stress response" “Regulation of hypoxia induced factor activity via alternative pre-IRNA splicing”
- International bureau of the BMBF, Germany, on „Bivalent Hantavirus vaccine for Europe based on recombinant proteins”
- Support for the integration of “newly associated states” (NAS) in the European research area. Project “Biotechnology Centre of Excellence” (BIOCEL).

Designed by IDEA ARTIS
Printed by SOLIDARITY

Institute of Biotechnology
Centre of Excellence

V.Graičiūno 8
02241 Vilnius-28
Lithuania
Phone: + 370 5 2602 103
Fax: + 370 5 2602 116
E-mail: office@ibt.lt
www.ibt.lt