Meeting report

A European focus on proteomics

Christian MT Spahn*, Hans Lehrach† and Peter R Jungblut‡

Addresses: *Institute of Medical Physics and Biophysics, Charité, Berlin, Germany. †Max Planck Institute for Molecular Genetics, Berlin, Germany. ‡Max Planck Institute for Infection Biology, Berlin, Germany.

Correspondence: Hans Lehrach. E-mail: lehrach@molgen.mpg.de

Published: 16 April 2004

Genome ***Biology*** 2004, **5:**322

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2004/5/5/322

© 2004 BioMed Central Ltd

A report on the First International Symposium of the Austrian Proteomics Platform, Seefeld, Austria, 26-29 January 2004.

The perfectly organized Austrian Proteomics Platform meeting that took place in the beautiful environment of the Austrian Alps heralded a major initiative of Austrian researchers and science-politicians in launching a program in proteomics. Many leading experts in functional proteomic studies gave an excellent overview of state-of-the-art proteomic studies that certainly encouraged the many young scientists participating at the meeting.

Bridging genomics and proteomics

Ronald Davis (Stanford Genome Technology Center, USA) presented a number of techniques for functional genomics, genotyping of single-nucleotide polymorphisms (SNPs), and detecting mutations. A particular focus of this work is the detailed analysis of gene function in yeast, using the complete set of gene-deletion mutants that is now available for this organism. Jason Ptacek (Yale University, New Haven, USA) described the systematic analysis of yeast proteins in liquid as well as in protein-chip format. A major part of this analysis deals with the systematic characterization of protein kinases, as well as proteins serving as substrates for protein kinases.

H.L. compared the cell to a neural network that computes the phenotype from the genotype, additionally taking the environment into account. He highlighted new functional genomics and proteomics technologies and discussed the need for new integrated types of database, such as the GenomeMatrix database [http://www.genome-matrix.org/] of the German Resource Center for Genome Research (RZPD), and for computer models if we are to understand data from high-throughput genomics/proteomics experiments.

Advances in technology and data interpretation

Most current studies of proteomes are based on mass spectrometry (MS). Efforts in improving four aspects of such studies are key for a more comprehensive study of the human proteome - which is based on about 30,000 genes resulting in more than 1,000,000 protein species. The first need is for improved accuracy and speed of mass detection; the second is improved separation of complex (protein) mixtures; the third need is methods for computationally refining the signals obtained; and the fourth is ways of integrating the flood of results in intelligent databases.

Barry Karger (Barnett Institute, Northeaster University, Boston, USA) gave an overview of several current MS techniques. He emphasized the complementarity between different techniques. In a comparative study of 51 ribosomal proteins, only 32 proteins were identified by both of the favored current techniques - matrix-assisted laser desorption/ionization (MALDI) MS and electrospray ionisation (ESI) MS - whereas 8 proteins were found only by ESI-MS and 11 only by MALDI-MS.

Major improvements in technology from 1996 to date are a 150-fold increase in speed, by enhanced instrumentation and lasers with higher frequency. Current strategies of data analysis focus on de-noising using matched filtration and wavelet transformation for signal processing. Important steps in using MS to investigate the proteomes of tumor tissues are efficient methods of obtaining sample. Karger presented an automated laser-capture microdissection process that allows him to obtain about 10,000 cells, or 1-4 μg of total protein, in one cup. Relative quantification using $^{16}O/^{18}O$ exchange methods during trypsination resulted in

several differentially regulated proteins in a breast cancer study. The high potential of Fourier-transformed ion cyclotron resonance-MS (FTICR-MS) became reality in an experiment in which 1 μg protein from 10,000 cells was used and 820 proteins were identified. New evaluation programs are in progress for the high mass-accuracy data obtained by FTICR-MS. Further improvements may be expected by miniaturization of columns down to 20 μm inner diameter and flow rates in the range of 20-50 nl/min, reaching a detection limit of 5-10 amol of peptide.

The need for improvements in data acquisition, data analysis and data assembly was also discussed by John Yates III (Scripps Research Institute, San Diego, USA), who in particular emphasized their importance for the analysis of complex protein mixtures in quantitative shotgun proteomics. Applying subtractive proteomics to nuclear envelopes, Yates and colleagues identified many novel envelope proteins that could be mapped to chromosomal regions linked to a variety of dystrophies. Given that metabolic labeling is helpful in studying the dynamics of proteomes, Yates reported progress towards metabolic labeling in tissue by feeding ^{15}N-labeled proteins to rats.

Isabel Feuerstein and Sandra Morandell gave a 'young investigator talk' jointly from the labs of Günther Bonn and Lukas Huber (all from the University of Innsbruck, Austria). Feuerstein pointed out that phosphoproteomics is important, since there are about 100,000 potential phosphorylation sites in the cell's proteome, but only 2,000 are currently known. She presented IDA-Fe^{3+}-cellulose, a newly developed cellulose sorbent based on immobilized imino-diazetic acid (IDA) ion exchange beads, for enhanced enrichment of phosphopeptides in immobilized metal affinity chromatography (IMAC). Morandell used this novel material for phosphoproteomics of the mitogen-activated protein (MAP) kinase signaling pathway.

Proteomics and disease

Several speakers described promising proteomics efforts that aim to identify new targets for the diagnosis or treatment of disease. Annemarie Poustka (Deutsche Krebsforschungszentrum (DKFZ), Heidelberg, Germany) presented a program of research that intends to understand the underlying principles of disease, especially tumor etiology and development. Part of the program is the sub-classification of cancers using gene-expression profiling, in order to find changes in gene expression between cancer cells and normal tissue. The project is based on arrays (chips) bearing 36,000 human cDNA clones for hybridization, and tissue-specific chips. By hierarchical clustering of genes expressed in kidney tumors, Poustka showed that there is a significant association with chromosomal aberrations for 10-20% of genes that are differentially expressed in tumors. Among the identified targets there was a significant upregulation of genes involved in glycolysis, whereas the genes involved in gluconeogenesis were downregulated. Genes showing clear expression differences in tumors are now being screened in high-throughput functional assays, to identify their function within the cell and in the hope of unraveling the network of molecular processes in different tumors.

Tom Conrads (National Cancer Institute, Frederick, USA) and John Semmes (Eastern Virginia Medical School, Norfolk, USA) spoke about cancer profiling using crude biological samples, such as serum. Conrads stressed the importance of translational cancer research, which can be applied to real clinical conditions; he investigates whether patterns of proteins or peptides in blood can be used for histopathology. Pointing out the power of modern FTICR-MS techniques he reported the first success with a controlled ovarian cancer dataset, for which he can distinguish tumor from control samples with 100% sensitivity and 100% specificity. A similar study presented by Semmes uses serum-protein fingerprinting in prostate cancer.

Denis Hochstrasser (Geneva University Hospital, Switzerland) argued that because of the proteome's complexity it is important to decide at which level to work, and that simplification might be a good strategy for asking the relevant questions, especially for smaller labs. As an example, Hochstrasser used the clinical situation for stroke patients. Since the worst outcome of stroke is death, he used the dead brain as a model and found fatty-acid-binding protein as a new marker for stroke. Moreover, he could differentiate between ischemic or hemorrhagic stroke using the apolipoprotein C-III gene as a marker. In a technological development, Hochstrasser described how a molecular scanner can be used for scanning two-dimensional electrophoresis gels for MS. In the same way, direct blotting of proteins from cryostatic sections from brain results in discovery of the position of proteins within the tissue. Absolute prerequisites for proteomic investigations are optimal documentation, controls and appropriate data-tracking and retrieval, so as to provide possibilities for data mining, for which laboratory information management systems (LIMSs) are required.

A role for proteomics in studying infectious diseases was presented by P.R.J. Subtractive analyses by two-dimensional gel electrophoresis plus MS, and by isotope-coded affinity tag liquid chromatography MS (ICAT-LC/MS), resulted in identification of about 30% of *Mycobacterium tuberculosis* predicted genes at the protein level, and led to vaccine candidates with promising results in preclinical vaccination experiments. A multiparameter selection of protein candidates, combining data from electrophoresis, MS, immunoproteomics, prediction of T-cell epitopes and others, also revealed vaccine candidates for *Helicobacter pylori*, which again have been tested in a mouse model successfully. An international proteome database, called Proteome 2D-PAGE

database [http://www.mpiib-berlin.mpg.de/2D-PAGE], is now open for the scientific community.

Hanno Langen (Roche Genetics, Basel, Switzerland) discussed proteomics in the pharmaceutical industry for the purpose of discovering and validating new molecular drug targets, which he named three-dimensional proteomics. In *Bacillus subtilis,* 1,505 different proteins were identified by peptide-mass fingerprinting and tandem MS experiments. Quantification is also possible by MS matching (finding hits) without isotope labeling. He also emphasized that in an industrial environment, a LIMS is a prerequisite for proteomics.

Exploring proteomes and protein complexes

According to Joel Vandekerckhove (Ghent University, Belgium), the dynamic range of proteins in a human cell ranges from one molecule to up to 10^{10} molecules per cell. This fact, together with difficult chemistry and no obvious amplification procedure (unlike nucleic acids), means that proteomics is a much more difficult challenge than genomics. Complementary methods are necessary to get access to a representative part of a proteome. One available method is combined fractional diagonal chromatography (COFRADIC), a peptide isolation procedure based on diagonal electrophoresis and diagonal chromatography.

The problem of characterizing the genomes of organisms with unsequenced genomes was addressed by Andrey Shevchenko (Max Planck Institute for Molecular Cell Biology and Genetics, Dresden, Germany). He pointed out that even homologous proteins are difficult to identify. As an example he gave human dihydrodiol dehydrogenase (DDH) and alligator alcohol dehydrogenase (ADH), which share 75% sequence identity but no identical peptide in MS. Possible ways to overcome this problem are the use of multiple tags in combination with better computer algorithms and the return of *de novo* sequencing.

Mathias Dreger (Free University of Berlin, Germany) introduced subcellular proteomics and stressed that the subcellular distribution of components is an important property of the proteome. He gave examples for the analysis of the nuclear envelope and the analysis of spinal cord synaptic membrane proteins, which are important for pain research. C.S. went one step further within subcellular organization and emphasized the increasing importance of protein complexes and macromolecular machines. He presented cryo-electron microscopy as a promising technique for structural proteomics of such complexes, and introduced the Ultra-Structure Network in Berlin, which is in the early stages of development, as a research initiative that aims to achieve high-throughput analysis of macromolecular complexes by MS and cryo-electron microscopy. Overall, the meeting was a genuine success and a promising start for GENAU, the Austrian functional genomics program, and especially its proteomics component.