Proteomics: key technology in drug discovery

The first IBC UK conference on *Technical Advances & Applications of Proteomics in Drug Discovery, Clinical Development, Safety Testing & Diagnosis* was held at the London Marriott Hotel on the 16–17 February 1998. The meeting focused on various aspects of proteomic technology and its application in drug discovery, diagnostics and the study of disease processes. The Proteome has been defined by Wilkins, M.R. *et al.* [*Biotech. Genet. Eng. Rev.* (1995) 13, 19–50] as the complete set of proteins encoded by the genome. Proteomics may be viewed as the global analysis of gene expression using a plethora of state-of-the-art techniques available to resolve (high resolution 2D gels), quantitate (phosphorimager, special scanners), identify and characterize proteins (microsequencing, mass spectrometry), as well as to store, communicate and interlink protein and DNA sequence and mapping information (bioinformatics). The major targets for drug discovery are proteins [Dr Leigh Anderson, Large Scale Biology (LSB), Rockville, MD, USA] and ~1,000 of the 100,000 believed to comprise the human proteome may represent targets (Dr Walter Blackstock, Glaxo Wellcome, Stevenage, UK). With this in mind, it seemed most timely to gather a broad spectrum of scientists to discuss recent developments in proteomic technology as well as strategies to improve the selection of putative targets.

Spotlight on the proteomics approach

The meeting was opened by Dr Anderson (LSB), a pioneer in the field of proteomics who, together with Dr N. Anderson, was amongst the first in the late 1970s to grasp the potential of 2D gels (developed independently by O'Farrell, P. and Klose, J.) to resolve and catalogue the vast repertoire of human proteins expressed in normal and disease conditions. Dr Anderson emphasized that mRNA and protein levels do not always correlate; hence, focusing on the proteins as drug targets has certain advantages compared with mRNA-based technology. In addition, proteomics may yield information concerning the half-life of proteins and will certainly provide valuable insight into their post-translational modifications. Several contributors commented on the advantages of proteomics over other approaches, although there was a clear consensus that an integration of proteomics, genomics, phage display technology and animal models would maximize the discovery of potential targets.

Dr Anderson also highlighted the usefulness and potential of the proteomic approach to identify quantitative changes in rat liver expression profiles associated with toxicity of drugs and other xenobiotics. The data, which are systematically stored in the Rodent Molecular Effects Database, are expected to yield important information about the molecular mechanisms of toxic responses. Similar studies were presented by Dr Sandra Steiner (Novartis, Basel, Switzerland) who, in collaboration with LSB (access to their databases) and Innogenetics (Ghent, Belgium), is applying proteomic tools to routine molecular toxicology studies using rat liver and kidney material.

Proteomics and the industry

Other presentations dealt with issues such as protein expression analysis in the context of drug discovery (Dr Wayne Bowen, Pharmagene, Royston, UK) and the potential of proteomics in preclinical and clinical development (Dr Chris Moyses, Oxford Glycosciences, Abingdon, UK). Dr Bowen underlined the importance of using human biopsy material in pre-clinical discovery and put forward Pharmagene's strategy for drug discovery which is based on the analysis of protein expression profiles – determined by dot blotting using a high-density grid of human proteins – of a selected family of key proteins (for example, G-protein coupled receptors, ion channels or components of signalling pathways) that are established candidates for intervention. Concentrating on a small number of proteins, for example multiprotein complexes, rather than on the whole proteome was also argued for by both Dr Mathias Mann [European Molecular Biology Laboratory (EMBL), Heidelberg, Germany, but now at Odense University, Denmark] and by Dr Blackstock. By contrast, Dr Moyses pointed towards the high cost involved in bringing a new drug to the market and emphasized the need to improve the identification, characterization and selection of prospective targets. With this in mind, Oxford Glycosciences has established the whole proteomic technology in-house and has developed software and fluorescence dyes for sensitive high-throughput analysis of the proteome both in health and disease. Their approach was illustrated by a recent study of primary hepatocellular carcinomas and is expected to pave the way for a forthcoming systematic survey of other cancers, such as colorectal and prostate cancers.

Dr Craig Wilde (Incyte Pharmaceuticals, Palo Alto, CA, USA) underlined Incyte's interest in the field of proteomics, by announcing their collaboration with Oxford Glycosciences. This is a powerful alliance, as the availability of complete genome sequences will help

Copyright © Elsevier Science Ltd. All rights reserved.
1359-6446/98/$19.00. PII: S1359-6446(98)01184-2

proteomics enormously. He also introduced Incyte's new platform for delivery of DNA data products and stressed the importance of improving DNA sequence data for reliable use of databases for proteomics. The last talk of the morning was delivered by Dr Matthew Davison (Zeneca Pharmaceuticals, Alderley Edge, UK) who carefully analysed all the problems associated with establishing a proteomics group in a pharmaceutical company.

Protein analysis, identification and characterization

Dr Mann has been a pioneer in the application of mass spectrometry (MS) for the sensitive identification and sequencing of proteins. He presented an elegant analysis of entire protein complexes (Arp2/3 complex, spliceosome) using a combination of affinity purification and a high accuracy mass peptide-mapping technique. Ultimately, he aims to integrate genomic data with the proteome in a fashion similar to that described by Prof. Julio E. Celis [Danish Centre for Human Genome Research (DCHGR), Aarhus, Denmark].

The next speaker (Dr Blackstock), supported the idea of concentrating proteomic research on a selected set of proteins, such as complexes and organelles (i.e. cell map model of proteomics), as he believes it would be very difficult to analyse quantitatively all of the protein changes observed in tissues and cells under various physiological conditions (i.e. surrogate model of proteomics). He briefly mentioned Glaxo Wellcome's involvement in identifying Golgi proteins, and like other speakers commented on the need to automate the gel technology and to develop better software for image analysis and image range.

The integration of innovative chromatographic separations with MALDI-TOF-MS was presented by Dr George J. Vella (Perseptive Biosystems, Framingham, MA, USA). This was an interesting talk, as many people would like to do without 2D gels, which are difficult to reproduce, even when using immobilized pH gradients. He mentioned the advantages of perfusion over beads, as well as the possibility of carrying out interactive protein analysis using immobilized and soluble proteins. As far as MS was concerned, Dr Vella underlined the need to automate some of the search routines and to reduce the error to improve the mass accuracy. In other words, quality data is required, a fact that was also stressed by both Dr Mann and Dr Martin Page (Oxford Glycosciences, Abingdon, UK). Like many others, Dr Vella envisaged modern drug discovery as using a plethora of techniques that included genomics, proteomics, high-throughput lead generation, combinatorial chemistry and functional genomics. The functional analysis of unknown genes was covered by Dr Stephen J. Fey [Centre for Proteome Analysis (CPA), Odense, Denmark], while Dr Robert M. Frederickson (University of Washington, Seattle, WA, USA) elaborated on novel 'hybrid technology', based on the yeast two-hybrid technique, to elucidate protein–protein interactions.

Application of proteomics

In general, the second day of the conference was more exciting as many of the contributors illustrated, with experimental data, the potential and limitations of proteomics in the quest for drug targets and clinical applications.

Clinical diagnostics

Dr Kevin Johnson (Cambridge Antibody Technology, Melbourn, UK), showed the potential of phage antibody libraries to produce specific antibodies for diagnostic purposes. In particular, the company's main interest lies in secreted proteins for their potential therapeutic value. The experimental approach consists of searching EST databases for sequences that may encode secretory proteins. Thereafter, they synthesize the peptides, screen the phage library using panning techniques and prepare the antibodies. Imagine the power of the technology if it were possible to screen 2D-gel blots containing thousands of proteins. This technology is still under development, but it represents a powerful addition to the sophisticated technology already available for protein analysis.

The application of 2D PAGE databases to the study of gene expression in cardiovascular disease, and after transplantation was presented by Dr Michael J. Dunn (Harefield Hospital, UK; see http://www.harefield.nthames.nhs.uk/nhli/protein). One of the aims of his group is to develop noninvasive methods for detecting acute rejection, a study that is being carried out in collaboration with the CPA in Odense.

Drug discovery

Dr Paul Moore from Human Genome Sciences (Rockville, MD, USA) described their efforts to discover novel therapeutic proteins using large-scale cDNA databases. Their approach to gene discovery is based on automatic high-throughput DNA sequencing and the preparation of numerous cDNA libraries from normal adult tissue, foetal tissue, diseased tissue, cell lines and differential/subtracted conditions (such as normal vs diseased conditions).Selection of candidate genes takes place by similarity/homology searches, motif analysis, phenotypic association (i.e. differential gene expression) and also molecular modelling. Several potential therapeutic proteins have been identified using this approach, some of which seem to have potential medical applications. The presentation was illustrated with examples of the myeloid progenitor inhibitor factor 1 and the keratinocyte growth factor 2 (involved in wound healing).

Dr Page returned to the problem of drug discovery and to the role that proteomics will play in accelerating the various steps involved – target identification, target validation, drug discovery (efficacy, selectivity and mode of action), *in vivo* properties and clinical trials. He concurred with others that the quality of the data is more important than the volume, because a mistake in choosing a target may prove costly. In addition, he illustrated the use of Oxford Glycoscience's proteome technology to study the effects of 5-fluorouracil and OGT-719 on the protein-expression profiles of the hepatoma cell line Huh7. Some protein, as well as clinical, data have

already been collected, representing what may be the beginning of a proteomic 2D PAGE database.

Bioinformatics

Prof. Celis continued with the theme of building 2D PAGE databases, which he illustrated with the use of the large human keratinocyte database in the study of bladder squamous cell carcinomas. These cancer cells share many properties with keratinocytes and, therefore, data stored in the large keratinocyte database (http://biobase.dk /cgi-bin/celis) has been instrumental to identify markers that may differentiate various histopathological types. Their studies have also led to the identification of a specific biomarker, psoriasin, that is externalized to the urine by SCC-bearing patients. He discussed the problems of using proteomic technology to study heterogenous tissues and tumours, and underlined the importance of building comprehensive databases that contain information of a critical mass. These points were also stressed by Dr Peter Mose Larsen (CPA), who presented a well-performed, quantitative analysis of the protein changes observed in pancreatic β-cells in the islets of Langerhans during the development of diabetes, both in rat and human. Their studies illustrated how difficult it is to pin-point protein changes that are directly associated with the disease.

The power of bioinformatics in proteomics was explored by Dr Dennis Hochstrasser (Geneva University Hospital, Switzerland); his group has been instrumental in stimulating scientists to publish their 2D PAGE data on the Internet. He outlined the protein search tools available at the SWISS-PROT database (http://expasy.hcuge.ch/sprot/sprot-top.html) and described how proteomic technologies, in particular 2D PAGE databases, are being used in his laboratory to study disease processes.

Overall

In general, the conference seemed to satisfy most peoples' expectations, in spite of the fact that in some cases only limited experimental data was disclosed. Nevertheless, proteomics is here to stay and academics must learn to live with the idea that most of the data generated by the pharmaceutical industry may not readily reach the scientific community.

Julio E. Celis
Danish Centre for
Human Genome Research
Build. 170, Ole Worms Allé
DK-8000 Aarhus C, Denmark
tel: +45 89 422 880
fax: +45 86 131 160
e-mail: jec@biokemi.au.dk