was defined as the bringing together of data, algorithms and interfaces to assist discovery scientists in their decisions about what to test and make next. Several applications illustrated the efficient use of a combination of standard tools, such as RS3, Diva, Tsar, MOE and Oracle.

Evolutionary chemistry

Evolutionary breeding methods for small molecules were introduced by Lutz Weber (Morphochem; http://www.morphochem.com). These methods have been integrated into a software called MolMind™, which makes use of heuristic evolutionary methods for synthesis planning. It works with Morphochem's proprietary database of multiple component reactions and available starting materials that have been implemented on their robotic systems. MolMind™ is used to discover new multicomponent reactions that enable the synthesis of natural-product-like molecules, as well as the design of compound libraries around any lead structure of interest.

Alexander Tropsha (University of North Carolina; http://www.unc.edu) questioned whether traditional QSAR-model quality techniques, such as leave-one-out cross-validated R^2, could be used as proper indicators of the predictive ability of the models. He proposed novel parameters that characterize the statistical significance of QSAR models in a more rigorous way and their use in screening chemical databases and virtual databases for potentially bioactive molecules.

Requirements for data integration in pharma and biotech R&D present highly complex challenges, as described by Herschel Weintraub (CADDinformatics). A wealth of data, including chemical structures, genome sequences, screening results, Laboratory Information Management Systems (LIMS) data, ADME/Tox data and cheminformatics data, is continuously being collected. These data must be accessible to the research scientists and management in a timely and efficient manner. A key issue is the integrity of the systems that contain validated data and approaches to undertake such a massive challenge to data integration were presented.

Concluding remarks

This conference was rich in interesting talks, highlighting novel ideas and useful views. Expressions such as 'unique proprietary technology' and 'paradigm shift' have been seldom heard and this gave much more weight and interest to all the new ideas that were presented and debated. Several speakers described unsatisfactory or unexpected results as well as successes, stressing the fact that a failure is not necessarily something negative. Indeed, more and more scientists now seem ready to share their experience, even when things do not work as expected. After several years of staggering claims, the age of reason (and truth) seems to have been reached.

New routes for drug discovery

Kewal K. Jain, Bläsiring 7, CH-4057 Basel, Switzerland; tel/fax: +41 61 692 4461, e-mail: jain@pharmabiotech.ch

The recent IBC conference *Proteomics: Delivering New Routes to Drug Discovery* (6–9 May 2002, Philadelphia, PA, USA) continued its tradition of providing a broad coverage of the application of innovative proteomic technologies for drug discovery and clinical applications. Selected presentations are described here briefly. Updated proteomic technologies are described elsewhere [1].

Beyond the genome to clinical proteomic diagnostics

In the keynote address, Emanuel Petricoin (FDA–NCI Clinical Proteomics Program; http://www.fda.gov) described the paradigm shift in molecular medicine as the choice of therapy by category of disease, molecular profiling of tissue cells and therapy tailored to individual patients.

Proteomics has an important role in these concepts as it is not a mere catalogue of proteins but represents cellular network and signal pathways. Cancer is considered to be a proteomic disease at the cellular level and cancer profiles are made from laser capture microdissection (LCM) of cancer tissue taken from patients. The FDA program uses low-profile proteomic technologies for diagnosis and drug discovery in cancer and have justified a prospective population-based assessment of proteomic pattern technology as a screening tool for all stages of ovarian cancer in high-risk and general populations [2]. A five-year follow-up of this study showed that the accuracy rate of diagnosis of carcinoma of cancer was 99%.

David Misek (University of Michigan; http://www.umich.edu) described the proteomic technologies for cancer marker identification with particular application to lung cancer. Major advantages of cancer marker identification are the ability to profile tumor tissue and uncover a variety of markers: overexpressed proteins, post-translationally modified

 1359-6446/02/$ – see front matter PII: S1359-6446(02)02354-1

proteins, translocated proteins and cleavage products. Markers were detectable in 90% of lung adenocarcinoma cases. The CARET feasibility study showed that anti-annexin antibodies could be detected in serum samples collected a year before the clinical diagnosis of lung cancer.

Mike Dunn (Kings College London; http://www.kcl.ac.uk) reviewed the current status of heart disease proteomics. MS analysis has identified a set of proteins in the heart that were altered with acute or chronic rejection, and ELISA assays have been used to identify proteins that are potential non-invasive markers of acute and/or chronic rejection.

Pharmacoproteomics

Pharmaceutical proteomics, using the ProGEx™ platform, which involves quantitative disease and drug characterization, was described by Sandra Steiner (Large Scale Biology Corporation, LSBC; http://www.lsbc.com). LSBC has studied the effects of >100 compounds in animal studies and demonstrated the use of protein expression data in SAR, mechanism classification and lead selection. LSBC also conducts an ongoing program examining the mechanisms (therapeutic and toxic) of marketed pharmaceuticals, of which >50 have been studied so far.

Jan van Oostrum (Novartis; http://www.novartis.com) used the example of Bengamide (a novel sponge-derived marine natural product) surrogate marker. In the genomic approach, when H1299 cells are exposed to Bengamide, there is no transcriptional response but in the proteomic approach, >20 proteins are expressed. H1299 cells are then cultured for 24 hours in the presence of marker LAF389. Matrix-assisted laser desorption ionization (MALDI)-MS spectra characterize the isoforms 14-3-3γ and targets for Bengamide-E and analogs are identified. Target protein expression is suppressed with methionine aminopeptidases (MetAp) siRNAs for validation of targets. Novel MetAps are identified by searching the Celera (http://www.celera.com) genomic database using sequence motifs and structural elements conserved in MetAp1 and MetAp2.

William A. Hanlon (Merck Research Laboratories; http://www.merck.com) described the use of proteomics to profile different compound chemical classes. The PPAR (peroxisome proliferator activated receptor) transcription factors, which regulate gene expression involved in lipid metabolism, were studied. A unique signature profile of the altered protein expression containing both overlapping and agonist selective components was obtained by the use of 2D-DIGE (differential gel electrophoresis) separation technology. Alteration in expression of several proteins unique to PPARα and PPARγ agonists were identified. Both of these receptors are potential targets for the treatment of dyslipidemia.

Novel techniques for protein expression profiling

Joachim Ostermann (Caprion Pharmaceuticals; http://www.caprion.com) presented Caprion's CellCarta Proteomics approach to identify the entire protein complement of organelles under normal and diseased conditions. Whereas the conventional approaches provide incomplete protein identification showing abundant proteins only, with no information on protein location, the CellCarta Cell Maps provide comprehensive protein identification including low abundance proteins with their location and orientation. Protein trafficking events are identified and there is reliable protein expression profiling. Mass intensity profiling system (MIPS) measures the intensity of peptide ions derived from a specific protein as a relative measure of protein abundance and generates a comprehensive expression profile for each sample. Advantages of MIPS include automation, comprehensive protein coverage and amenability to 'gel-free' analyses.

William Hancock (ThermoFinnigan; http://www.thermoquest.com) described new approaches based on LC–MS for the characterization of the proteome of complex biological samples. For example, blood plasma could contain a complete set of secreted cellular proteins and can be termed a secretome. Human plasma analysis by LC–MS can reveal markers of disease process; one example of this is the detection of low levels of hGH (human growth hormone) in the plasma of athletes where conventional assays for GH do not detect extraneously administered GH.

Celera's Scott Patterson described the acceleration of drug discovery through chromatography-based proteomics (Proteomics Factory) with simpler automation and higher throughput. This approach shifts the high-resolution separation from front-end (2D electrophoresis) to back end (LC–MS). Proteins are kept in solution so that a higher percentage of the sample is analyzed using affinity chromatography with access to the entire genome of the organism under study. Protein mixtures are digested for analysis at peptide level and significant in-house computational resources are required for this analysis. This high-efficiency proteomics technology is applied to small-molecule drug discovery [3].

Lead optimization by chemical microarrays

Chemical genomics implies the study of how small molecules interact with cells; the term chemoproteomics can be used when proteomic approaches are used. Holger Ottleben (Graffinity; http://www.graffinity.com) described chemical microarrays in which small molecules are immobilized on a carrier surface and provided in a screening-ready, high-quality standardized format. Incubation of these compounds with target protein yields comprehensive affinity fingerprints in a label- and assay-free procedure. The proprietary low molecular

weight compound libraries for chemical microarrays are based on small drug fragments and enable an insight into structural features early in the discovery process, thus providing the basis for rapid lead optimization. Large-scale fingerprinting enables one protein versus 50,000 chemicals to be screened per day. Genome-wide screening provides drugability indication, target prioritization and protein affinity annotation of small molecules.

Integrated proteomics

Several proteomic technologies are being integrated for a better understanding of drug targets and more effective drug discovery. Jeremy Nicholson (Imperial College of Science, Technology and Medicine, London, UK; http://www.ic.ac.uk) reviewed metabonomics: a systems approach to investigate the metabolic consequences of drug exposure, disease processes and genetic modification. Although several spectroscopic methods have been used, NMR spectroscopy is considered to be one of the most powerful methods for generating multivariate metabolic data [4]. NMR-based systems approach is used for drug toxicity screening to aid lead compound selection. Metabolic phenotyping (metabotyping) is also used for investigating the metabolic effects of genetic modification and modeling of human disease processes. One deliberate gene knockout can produce several metabolic disturbances. Metabonomics can thus be used as a functional genomics tool with applications in various stages of drug discovery and development.

Concluding remarks

Review of the conference highlights indicates the rapid advances in proteomics in the past year. Proteomic technologies are used increasingly in the drug discovery process and improvements in understanding the disease process and improved diagnostics through application of proteomics is contributing to the development of better medicines.

References

1 Jain, K.K. (2002) *Proteomics: Technologies and Commercial Opportunities.* (8th edn), Jain PharmaBiotech Publications, Basel, Switzerland

2 Petricoin, E.F. *et al.* (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359,572–577

3 Ryan, T.E. and Patterson, S. (2002) Proteomics in drug target discovery – high-throughput meets high-efficiency. *Drug Discov. World* 3, 43–52

4 Nicholson, J.K. *et al.* (2002) Metabonomics: a platform for studying drug toxicity and gene function. *Nature Drug Discov.* 1, 153–161