HUPO 3rd Annual World Congress on Proteomics

25–27th October, 2004, Beijing, China

Gilbert S Omenn[1] and Peipei Ping

[1]University of Michigan, 1150 W Medical Center Dr., Room A520C, Ann Arbor, MI 48109-0656, USA; Tel: +1 734 763 7583; Fax: +1 734 647 8148; gomenn@umich.edu

Expert Rev. Proteomics 2(1), 9–12 (2005)

The Human Proteome Organization (HUPO) held its 3rd Annual Congress in Beijing, China on 25–27 October, 2004, preceded by a weekend of workshops on HUPO Initiatives and Educational Sessions. Congress President Fuchu He (China National Center of Biomedical Analysis, China), Chinese officials, and Congress Co-chairs Sam Hanash (Fred Hutchinson, Cancer Research Center and University of Michigan, USA) and HUPO President John Bergeron (McGill University, Canada) welcomed more than 1600 participants. The scientific program committee, cochaired by Zhu Chen (Chinese Academy of Sciences, China) and Peipei Ping (University of California, CA, USA), arranged a splendid and exciting program with seven keynote addresses, 180 formal presentations in sets of 16 parallel sessions, and 500 poster presentations.

Among the keynotes, Matthias Mann (University of Southern Denmark, Denmark) presented advances in large-scale proteomics of organelles and signaling pathways. A marked increase in sensitivity and precision has been achieved with a protocol that employs nano liquid chromatography (LC) electrospray ionization (ESI) linear ion trap/Fourier transform ion cyclotron resonance (FTICR) mass spectrometry (MS) with a third MS analysis of fragments. This yields mass accuracy to less than 3 parts per million with high throughput. The third MS is performed while the Fourier transform obtains high-quality analytical spectra. Mann described protein catalogs for mitochondria, illustrated with cytochrome c oxidase deficiency in Leigh syndrome, and distinctive proteins in mitochondria of different tissues, with 50% similarity. More than 600 proteins were identified in centrosomes. From eight subcellular fractions of mouse liver, 1400 proteins were identified, described as "functional genomics at subcellular resolution", providing molecular detail for Ramon y Cajal's classic drawings of 100 years ago. John Yates (The Scripps Research Institute, USA) noted that data-dependent acquisition of spectra is a sampling process, "like daily polling". Yates described a suite of software for data-independent multidimensional protein identification technology (MudPIT) quantitation of identified peptides, using stable isotopes for metabolic labeling. One application is identification of 271 sumoylated proteins in yeast in localized multiprotein complexes. Yates also described enrichment of nuclear membrane proteins by subtractive analysis of endoplasmic reticulum and mitochondrial proteins. Leroy Hood (Institute for Systems Biology, USA) presented the broadest view in the context of regulatory gene networks and systems biology. This was illustrated by redirection of the development of sea urchins and disease-related perturbations in cancers. Hood presented new prostate cancer biomarker candidates from isotope-coded affinity tags (ICAT) and high-throughput, super-sensitive (<40 transcripts/million) gene expression molecular signature discovery methods. His collaborators interrogate individual cells with nanofluidics, nanowires and nanocantilevers. Hood expects us to have highly multiplexed disease screening methods that are predictive, preventive and personalized within 10 years. Zhu Chen demonstrated how chemical genomic approaches reveal gene expression networks underlying all-*trans*-retinoic acid-induced differentiation of acute promyelocytic leukemia (APL), arsenic trioxide (ATO)-induced apoptosis of APL cells, and ST1571/ATO-induced apoptosis of chronic myelogenous leukemia cells. Other keynotes included the lectures presented by Lee Hartwell (Fred Hutchinson Cancer Research Center, USA), Christian Brechot (French Institute of Health and Medical Research, France) and Chih-chen Wang (Chinese Academy of Sciences, China).

Parallel session headers included MS, nanotechnology, protein microarrays, affinity reagents, protein chemistry and protein structure; plasma, liver, functional, glyco-, neuro-, microbial and plant, disease, pharmaceutical and structural proteomics; and protein informatics. Affinity reagents is clearly a rapidly emerging field for translating discoveries from MS to higher throughput clinical and biologic applications. Matthias Uhlen (Karolinska Institute, Sweden) summarized the variety of methods for monoclonal, monospecific polyclonal and polyclonal antibody production in the mouse, rabbit, goat, sheep and chicken, using recombinant and natural proteins. This was achieved with a wide range of detection techniques from sandwich secondary antibodies to phage display to surface plasmon resonance. Qihong Sun (China National Center of Biomedical Analysis, China) highlighted the work in the Human Liver Proteome Project (HLPP). Larry Gold (Somalogic Inc., USA) described aptamer affinity and adaptability, and the striking power of new photoaptamer-based high-density arrays. These capture agents are single-stranded

10.1586/14789450.2.1.9

© 2005 Future Drugs Ltd

DNA sequences with a BrdU in place of thymidine so covalently bind to electron-rich amino acid side chains of protein analytes with high signal-to-noise ratios.

Naoyuki Taniguchi (Osaka University, Japan), Hui Zhang (Institute for Systems Biology, USA) and others described multiple methods for dissecting the chemistry and enzymology of glycoproteins, using glycospecific reagents to pull out this subproteome, and development of gene and protein databases.

Deletion of the core fucose from immunoglobulin (Ig)G molecules markedly enhances antibody-dependent cellular cytotoxicity, which may be useful for antibody therapies. S-L. Wu (USA) described a new platform for characterization of proteins with post-translational modifications (PTMs) (phospho- and glycosenaitivity increased by tenfold or more) by large peptide fragment analysis with a hybrid ion trap FT mass spectrometer. The aim is to detect subtle changes in glycosylation without altering protein expression level. Chris Sutton (PerkinElmer Life & Analytical Sciences, USA) and Koichi Tanaka (Shimadzu Corp. Kyoto, Japan) used methods similar to those of Mann in a matrix-assisted laser desorption/ionization quadrupole ion trap time-of-flight mass spectrometer to identify site-specific glycosylation moieties and phosphorylation status.

Hepatitis B and C infections and subsequent hepatocellular carcinoma (HCC) are very common throughout the world, particularly in Asia and Africa. Laura Beretta (Fred Hutchinson, USA) and Christian Brechot (keynote speaker, Pasteur/Necker Institute, France) were prominent among presenters in this field. Beretta reported three dozen proteins in lipid-enriched vesicles (lipid rafts) isolated from hepatitis C virus-replicating cell lines, and proteolytic fragments of certain proteins released into serum in HCC patients. Brechot focused on the pathways that lead to chronic inflammation, cirrhosis, and HCC; hepatitis B and C virus insert in hot spots affecting telomerase and calcium regulation (hepatitis B virus) or transforming growth factor-β-dependent signaling (hepatitis C virus).

Informatics is essential to proteomics. Rolf Apweiler (European Bioinformatics Institute, UK) stated a goal to transform 'Publish and Vanish' publications into valuable data resources through controlled vocabularies, careful and timely curation, and proteomics standards. Others have spoken about transforming databases into "knowledge bases". Henning Hermjakob (European Bioinformatics Institute, UK) described the HUPO Proteomics Standards Initiative (PSI) Molecular Interaction format and various tools, including IntAct and Cytoscape. Phil Jones (EBI, UK) presented a Distributed Annotation Server to query and publish sequence and positional features from UniProt via various accession numbers. UniProt went online in December 2003, joining the information contained in Swiss-Prot, TrEMBL and PIR (Manuela Pruss, EBI, UK). Ursula Hinz (Swiss Institute of Bioinformatics, Switzerland) described version 43.6 of Swiss-Prot, with emphasis on polymorphisms in health and disease, isoforms from alternative splicing and PTMs. InterPro (Sandra Orchard, EBI, UK) allows users to assign an unknown protein to a family and identifies domains, motifs, active sites, binding sites, and potential sites of PTMs within its sequence, integrating multiple other databases. Haipeng Wang and coworkers (China) reported a new tool, pLabel, for mining more information automatically from tandem MS spectra to overcome high false-positive rates of peptide and protein identifications. The same group in Shanghai presented a Fragment ion Formula Prediction method (Jingfen Zhang and coworkers). Eriksson and Fenyo (Sweden) demonstrated simulations of their protein identification error rate estimation. Yixue Li (Chinese Academy of Sciences, China) described a comprehensive software system for storing, sharing, and exchanging experimental data. Steven Brenner (University of California, Berkeley, USA) described phylogenomic reconstruction algorithms from evolutionary models, which should help identify proteins with conserved domains but not conserved sequences.

First prize in the poster category was awarded to Siu Kwan Sze of the Genome Institute of Singapore for a presentation on *Probing Protein Structure by Laser-Induced Oxidation and Mass Spectrometry* (27.14). An ultra-fast pulse ultraviolet laser produces extensive protein surface oxidation in nanosecond time scale, with generation of free hydroxyl radicals that react with amino acid residues and increase the mass by 16 Da, mapping the solvent accessible surface. Oxidized sites in several well-known proteins compare favorably with results from laborious x-ray crystallography and nuclear magnetic resonance (NMR) methods. Approximately 30 others received second or third place awards. Among so many presentations, the authors wish to highlight a poster by H-J Warnatz (Max-Planck-Institute for Medical Research, Germany) involving systematic proteomics to identify the protein products, their cellular localization and protein–protein interactions from the 285 genes mapped to human chromosome 21, of which 135 have no Gene Ontology annotation. Surely, chromosome-by-chromosome dissection of the human genome and proteome will become a major approach of broad interest.

HUPO is accelerating the development of proteomics through its initiatives, chaired overall by Sam Hanash. A total of 2 days of workshops and plenary sessions from each of the initiatives preceded the congress itself. These were extremely well attended and generated lots of useful discussion, as well as contributing many oral and poster presentations during the congress.

The HLPP, led by Fuchu He, Christian Brechot and John Bergeron, held center stage, with major investments to date by the Chinese and Korean governments. The HLPP has subgroups on sample collection and banking, proteome expression profiling, protein–protein interactions and open reading frame (ORF)eome, protein localization, PTM profiling, antibody banking, bioinformatics, rat and mouse organelle-level analysis, and proteomics of liver diseases. Extensive protein identifications from

fetal and adult human liver specimens were reported. Laura Beretta was made a cochair, succeeding Bergeron.

The Human Plasma Proteome Project (HPPP), led by Gilbert Omenn (University of Michigan, USA), has nearly completed its pilot phase, with a special issue of *Proteomics* scheduled for early 2005. The PPP drew upon more than 30 laboratories worldwide to analyze PPP reference specimens of serum, EDTA-, heparin- and citrate-plasma, and compare the performance of a great variety of depletion, fractionation and analytical platforms. From tandem MS or FTICR-MS (18 laboratories), 15,710 International Protein Index (IPI) (version 2.21) nonduplicated proteins were reported. These were consolidated to 9504 proteins with an integration algorithm that chose the best single representative for all proteins (including members of a protein family) perfectly matching the one or more peptide sequences. The 9504 proteins were identified by one or more laboratories; a requirement for confirmation of the protein identity by a second laboratory led to the core dataset of the PPP, with 3020 proteins. Many annotation methods are being applied. The various reference specimens gave relatively similar numbers of protein identifications, with a lot of variation for individual proteins (and, of course, coagulation-related proteins). There was extensive debate about the specificity of depletion methods and the relative virtues and limitations of fractionation, MS and search engine methods. Separate analyses examine immunoassay and microarray quantitation (and relation of concentration to detectability by proteomics methods) and direct surface-enhanced laser desorption/ionization MS findings (eight laboratories). The plasma has always been planned to be a final common pathway from the HLPP and Human Brain Proteome Project (HBPP).

The BPP, led by Helmut Meyer (Ruhr-University Bochum, Germany) with 18 participating laboratories worldwide, is aimed at understanding neurodegenerative disorders and aging. The BPP has chosen parallel initial studies with normal mouse brain (three age ranges) and human brain specimens from patients with Alzheimer's disease, Down syndrome and Parkinson's disease. Production of large numbers of antibodies against brain proteins is a key element, along with links to the PSI and PPP. ProteinScape has been installed on servers at individual laboratories and at the BPP center, using Minimum Information About a Proteomics Experiment (MIAPE) and mzDATA.

The HUPO PSI, led by Rolf Apweiler, has continued to expand its IPI database, which consolidates Swiss-Prot, RefSeq and The National Center for Biotechnology Information, with extensive curation and regular updates. The PSI has led consensus-building exercises that generated the MIAPE minimum reporting requirement for proteomics. These include:

- PSI Object Model: with controlled vocabularies and coordination with gene expression and other databases
- PSI-MS: for MS results, including recently released mzDATA (peaklists, peptide sequences and instrument parameters), collaborations with instrument vendors, and mzIDENT, which captures search engine output (due April 2005) in XML schema
- PSI-MI: for Molecular Interaction Data Exchange, with PSI-MI XML schema
- PRIDE: the Protein Identifications Database, which will become PSI-compatible in early 2005

The European Bioinformatics Institute/HUPO PSI serves as the repository and integrator for other initiatives from its close involvement in the HPPP, HLPP and HBPP initiatives, each of which has its own Data Collection Center connected with European Bioinformatics Institute.

The HUPO Human Antibody Initiative, led by Matthias Uhlen (Karolinska Institute, Sweden), is embracing the Sanger Institute protein atlas project in the UK, the Human Proteome Resource program in Sweden, several German antibody projects, antibody banking in the HLPP, structural genomics at RIKEN in Japan, and commercial developments in multiple companies.

The Education Program organized by Angelika Gorg (Technical University of Munich, Germany), Richard Simpson (Ludwig Institute for Cancer Research, Australia), Rong Wang (Mount Sinai School of Medicine, USA) and Peter Roepstorff (University of Southern Denmark) drew audiences of over 400 to duplicate sessions on protein separation, protein identification, quantitative proteomics and PTM proteomics. There is clearly a thirst, especially among younger scientists, for more knowledge and skills in this highly interdisciplinary field.

Meanwhile, the exhibit halls were packed with registrants. Exhibitors, from the European Bioinformatics Institute and the American Association for the Advancement of Science (publishers of *Science* magazine and *Science Online*) to instrument and reagent manufacturers reported hectic and productive interactions and high attendance at company-sponsored presentations and demonstrations.

It was clearly timely for China to host this congress. The number of scientists engaged in proteomics in China (and in Korea) is quite impressive. The investment in facilities and research programs, including the Chinese part of the HUPO LPP, is tremendous. Fuchu He led a group visit to the Beijing (National) Proteome Research Center, due to open in May 2005, and the site for a National Proteome Engineering Center in Zhongguancun Science Park, a huge, beautifully landscaped site with a Biochip Center and several pharmaceutical/biotechnology company facilities already open and two big hospital complexes (which will serve the Beijing Olympics in 2008) included in the site plan.

HUPO Awards for Distinguished Achievement in Proteomics were presented to Angelika Gorg for her decades of work on gel electrophoresis and to Rolf Apweiler for his leadership of the HUPO PSI and development of proteomics informatics at the European Bioinformatics Institute. Young-Ki Paik (Yonsei University, Korea), Secretary-General of HUPO, received the Distinguished Service Award for his ongoing

role in building the organization and its website. Also, for the first time, HUPO made 132 Young Scientist Awards.

Angelika Gorg and Matthias Mann announced plans for the 4th HUPO World Congress, in Munich, August 27th – September 1st, 2005.

Affiliations

- *Gilbert S Omenn, MD, PhD*
  University of Michigan, 1150 W Medical Center Dr., Room A520C, Ann Arbor, MI 48109-0656, USA
  Tel.: +1 734 763 7585
  Fax: +1 734 647 8148
  gomenn@umich.edu

- *Peipei Ping, PhD*
  Department of Physiology Medicine & the Division of Cardiology, David Geffen School of Medicine, Rm. 1619 MRL Bldg, University of California, Los Angeles, CA 90095, USA
  Tel.: +1 310 267 5624
  Fax: +1 310 267 5623
  pping@mednet.ucla.edu