CANCER GENOMICS & PROTEOMICS *2*: 159-198 (2005)

# ABSTRACTS OF THE FIRST INTERNATIONAL CONFERENCE OF THE HELLENIC PROTEOMICS SOCIETY

*Foundation for Biomedical Research of the Academy of Athens*

*May 22-25, 2005*

***Organizing Committee***

M. Fountoulakis, *Athens, Greece*
S. Kossida, *Athens, Greece*
H. Langen, *Basel, Switzerland*
P.G. *Righetti, Verona, Italy*
G.T. *Tsangaris, Athens, Greece*
A. *Vlahou, Athens, Greece*

## Sponsors

Amersham Biosciences, London, England
Antisel S.A., Thessaloniki, Greece
Applied Biosystems GmbH, Darmstadt, Germany
Beckman Coulter Inc, Fullerton, CA, USA
Biorad Laboratories S.r.l., Life Science Group, Milano, Italy
Biosolutions LTD, Athens, Greece
BioSURE – R&T Cell Co., Athens, Greece
Bruker Daltonik GmbH, Bremen, Germany
G. Kordopatis LTD, Athens, Greece
Interlab LTD, Athens, Greece
J.M. Kyriakidis SA, Athens, Greece
Malva SA, Athens, Greece
PerkinElmer, Life & Analytical Sciences, Milan, Italy
F. Hoffmann-La Roche LTD, Basel, Switzerland
Sigma-Aldrich Hellas, Athens, Greece
Thermo Electron Corporation, London, England
Varelas S.A., Athens Greece

**1**

## THE EVOLUTION OF FOCUSING TECHNIQUES: THE MARCH OF PENNIES, THE MARCH OF DIMES

Pier Giorgio Righetti

University of Verona, Laboratory of Proteome Analysis, Strada Le Grazie No. 15, Verona 37134, Italy

The present lecture on the evolution of isoelectric focusing (IEF) and immobilized pH gradients (IPG) is meant to dissipate the "defocusing of the origin of ideas" that, according to Kolin (1977) "*consists of blurring of historical perspective and usually takes the form of shifting the responsibility for discoveries and ideas to (preferably multiple) sources in the remote past*". The fog on the origin of ideas is usually accompanied by haze on the chemistry underlying the IEF/IPG principles, by which, *e.g.*, even in famous text books, one learns that the Immobiline chemicals are "amphoteric compounds", an odd notion indeed. This chronicle starts in 1961, when Svensson had the remarkable idea of focusing macroions in a pH gradient created by a multitude of amphoteric buffers, that went down to history as "carrier ampholytes" (CA), and ends in present days, with the remarkable performance of two-dimensional (2D) maps generated by coupling IPGs, in the first dimension, with SDS-PAGE (sodium dodecyl sulphate, polyacrylamide gel electrophoresis) in the second, orthogonal dimension. Highlights of the present lecture include:

- Who shipped who in the race for IEF paternity? H. Svensson Rilbe, A. Kolin or W.G. Kauman?
- The remarkable synthesis of CA buffers, a great deed in organic chemistry performed by a medical doctor (O. Vesterberg).
- The conversion of IEF from a preparative technique in sucrose density gradients to an analytical version performed in tiny polyacrylamide sausages and then in gel slabs (thin and ultrathin).
- The birth of IPGs, a unique event unveiled here, in the glorious town in Athens, in 1982.
- The long march of IPGs from narrow (max 1-pH-unit) to wider and wider gradients, encompassing just about the entire workable pH interval (pH 2.5 to 12). "Unlike chairman Mao's long march, that forced the red rebels, in 1934, to escape for a whole year the encircling and destruction from Chiang Kai-Shek troops, this march lasted some 10-12 years".
- The spreading of IPGs as a first dimension of 2D maps, with its unique advantages, such as reproducibility of spot position in the 2D plane and much sharper resolution.
- The evolution of IPGs from continuous gradients into membranes (single pI values) and amphoteric beads, most useful for pre-fractionation and for exploring the "deep proteome".

Righetti PG: Electrophoresis *25*: 2111-2127, 2004.
Righetti PG, Castagna A, Antonioli P and Boschetti E: Electrophoresis *26*: 297-319, 2005.

**2**

## PROTEOMICS TOOLS IN THE SEARCH, VALIDATION AND APPLICATION OF NEW BIOMARKERS

Hanno Langen

Roche Center of Medical Genomics, Switzerland

The major application of proteomics is the finding of new disease marker proteins, or so-called biomarkers. Biomarkers can also be used for patient stratification, *e.g.* in response prediction, to predict unwanted side-effects, as so called toxicology markers, and also to monitor disease progression or the positive effects of drug treatments.

Proteomics is traditionally a hypothesis-free approach to identifying new proteins, which were not known to be involved in disease processes. The proteins are separated by chromatography and electrophoresis and finally identified by various mass spectrometry techniques.

The main goal in the search for biomarkers is to identify these disease-specific proteins in body fluids, mainly plasma or serum. Many groups have tried but failed to identify such proteins directly in plasma by standard proteomics techniques. The main reason for this failure is the extensive dynamic range of protein in plasma. The main protein, albumin, is present at 50 mg/ml, whereas cytokines are present at only a few pg/ml. This means that there is a difference in protein abundance of 10 orders of magnitude. Even the best hypothesis-free fractionation and identification procedures based on mass spectrometry are not able to span this huge dynamic expression range. Therefore, we started our search for new biomarkers in disease tissue by proteomics and genomics with the assumption that tissue damage and remodeling leads to a leakage of the relevant proteins into plasma. We have established techniques to verify those markers in plasma based on antibody technologies. Examples from various diseases are given.

**3**

## THE HUPO BRAIN PROTEOME PROJECT

Helmut Meyer

Ruhr-University of Bochum, Germany

The HUPO Brain Proteome Project was initiated by Helmut E. Meyer, Bochum and Joachim Klose, Berlin, Germany, in 2003. It was a study of neurodegenerative diseases with the help of proteomic technologies. After a consolidation phase, the year 2004 was dedicated to two

pilot studies: i) The proteome and transcriptome analysis of mouse brain samples derived from 3 different age stages (embryonic, juvenile and adult, highly standardized) to obtain a reliable reference database as well as to examine age-related changes. ii) The proteome analysis of human brain samples derived from autopsy and biopsy, respectively, to obtain a reliable reference database as well as to study post-mortem protein stability.

Groups from Belgium, China, Germany, Greece, South Korea, Ireland, Switzerland, the UK and the US agreed to analyze these samples and to use the software ProteinScape (Bruker/Protagen, Germany) as a common, standardized proteomics database platform. A Data Collection Center (DCC) has been implemented at the Medical Proteom-Center, Bochum, and has already received the first data sets of the participants. Independent colleagues will start a standardized re-analysis of the data in April 2005, resulting in highly confidential gene and protein lists as well as new insights into proteomics analysis strategies. The first results and data will be presented.

Based on this knowledge, further studies will be designed, starting with a mouse model workshop on neuro-degenerative diseases on June 1, 2005, in Doorwerth, The Netherlands.

**4**

## HEAT SHOCK PROTEIN 27 IS ASSOCIATED WITH FREEDOM FROM TRANSPLANT-ASSOCIATED CORONARY ARTERY DISEASE FOLLOWING HUMAN CARDIAC TRANSPLANTATION

Michael J. Dunn[1], Ayesha I. De Souza[2] and Marlene L. Rose[2]

[1]Proteome Research Centre, Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland;
[2]Transplant Immunology, National Heart and Lung Institute, Imperial College London, Harefield Hospital, Harefield, Middlesex, U.K.

Transplant-associated coronary artery disease (Tx-CAD) is the major long-term complication following heart transplantation. We used 2-DE to compare protein profiles of early biopsies (<21 days) and late biopsies (>9 years) from the same patient to identify proteins that may be protective against Tx-CAD. A total of 141 proteins were significantly changed between the groups and 81 of these have been identified by LC-ESI-MS/MS. The greatest change (>20-fold) involved spot 3306, which was identified by MS/MS as Hsp27. This was strongly expressed in late biopsies taken from patients without disease, but was absent from late biopsies with Tx-CAD. The observed p*I* of protein spot 3306 was 5.42, whereas the theoretical p*I* of Hsp27 is 5.98, suggesting that the protein is a phosphorylated form of Hsp27. The phosphoprotein-specific stain, Pro-Q Diamond, confirmed Hsp27 spot 3306 to be phosphorylated. Hsp27 has three known phosphorylation sites at Ser-15, Ser-78 and Ser-82. MS/MS analysis of spot 3306 provided unequivocal evidence for phosphorylation of Ser-82 and indicated that Ser-15 is not phosphorylated. It was not possible to verify the phosphorylation state of Ser-78, as the tryptic peptide which includes this residue is too small to be detected by LC-MS/MS. However, Western immunoblotting with antibodies specific for the phosphorylated forms of Hsp27 demonstrated protein spot 3306 to be phosphorylated at serine residues 78 and 82, thus confirming protein spot 3306 to be diphosphorylated. Immunohistochemical analysis of tissue biopsies validated that Hsp27 was more abundantly expressed on biopsies free of Tx-CAD and, moreover, showed it to be localised to blood vessels. In contrast, vessels from patients with Tx-CAD did not express Hsp27. In conclusion, our results demonstrate that vascular expression of diphosphorylated Hsp27 is associated with freedom from vascular disease after cardiac transplantation. Understanding the mechanism of this protective effect may provide new opportunities for therapeutic intervention.

**5**

## A PROTEOMIC APPROACH TO DISCOVER RELEVANT MARKERS TO DIAGNOSE BREAST CANCER

Raymonde Joubert-Caron[1], Ludovic Canelle[1,2], Julie Hardouin[1] and Michel Caron[1]

[1]Laboratoire de Biochimie des Protéines et Protéomique (LBPP), Université Paris 13, UFR SMBH Léonard de Vinci, 93017-Bobigny, Cedex;
[2]R&D Immunoessais et Protéomique, bioMérieux, Chemin de l'Orme, 69280 Marcy l'Etoile, France

Tumor markers are substances that can be detected in higher-than-normal amounts in the body fluids of some patients with certain types of cancer. Markers are produced either by the tumor itself or by the body in response to the presence of cancer (*i.e.* autoantibodies). Autoantibody detection may be informative for the early detection and diagnosis of some types of cancer. In that context, proteomics provides major opportunities for screening and identifying autoantigens.

We followed a strategy based on a "serological proteome analysis" (SERPA) combining two-dimensional (2-D) electrophoresis, immunoblotting and image analysis. Parameters relevant to overcoming the variation of protein profiles prob on the 2D-blot were studied. Matching by image analysis permits the exact localization of particular

protein spots in preparative gels for further identification by mass spectrometry (MS). A set of 40 2D-blots was probed with 20 sera from patients with breast cancer (BPC) and 20 sera from healthy volunteers. In the protein profiles submitted to immunodetection, 15 proteins identified by MS were immunodetected in both BCP and healthy people. Seven spots reacted preferentially with BPC sera. Of these, two spots were not identified at all by MS or by MS/MS. One spot was identified as the enolase alpha subunit (P06733), with an incidence of 80% in the BPC group compared to 50% in the control group. Moreover, four spots were preferentially detected in BPC sera. MS or MS/MS identified these spots as various isoforms of three proteins: glucose-6-phosphate 1-dehydrogenase (P11413), heat shock 70 kDa protein1 (P08107) and dihydrolipoyl dehydrogenase (P09622). These proteins are widely expressed in normal cells. However, it is noticeable that the isoforms recognized preferentially by BPC differed from those of the same proteins recognized by almost all sera tested. The preferential immunodetection of specific isoforms of G6PD, HS71 and DLHD suggests that post-translational modifications are responsible for the appearance of autoantibodies

**6**
**HUMAN PROTEOME AND CANCER RESEARCH**

M. A. Molinaro

U.O. Radiotherapy and Radiobiology Hospital "Pugliese-Ciaccio", Catanzaro, Italy

Never before in the history of the fight against cancer have we enjoyed such a dramatic increase in our fundamental knowledge of the molecular mechanisms underlying neoplastic growth. This has led to the identification of new therapeutic options, with the promise of further important developments in the near future. The combination of genomics and proteomics approaches is the new tool in cancer research and possibly in the setup of new systems to evaluate patients and for individual therapeutic protocols. This new molecular medicine, based on proteome science, will allow discrimination between disease subtypes that are not recognizable using traditional pathological criteria. Moreover, specific genetic events, involved in cancer progression, may be identified and understood. Cell functioning is mainly related to proteins, and there are many protein modifications, like differential RNA splicing and post-translational modifications, not detectable from nucleic acid sequences. Furthermore, the study of protein interactions is broadening the vistas of cell-life regulation knowledge. For example PTEN, a tumor suppressor, functions as a regulator of both cell cycle progression and apoptosis. The PTEN gene is frequently mutated or deleted in several malignancies including human hepatocellular carcinoma (HCC). The clinical significance and prognostic value of PTEN expression in HCC remain obscure. Using immunohistochemistry, the PTEN protein expression in 46 tissue sections, collected from surgically resected HCC patients' was analysed. A significant positive correlation was identified between low PTEN expression in the HCC and increased expression of iNOS and COX II in the surrounding liver. The overall survival was significantly longer for the HCC patients with high PTEN expression than patients with low PTEN expression. Univariate analysis revealed PTEN expression as an independent prognostic factor for patient survival. The study demonstrated that down-regulation of PTEN in the tumor is an important step in HCV-positive cirrhotic hepatocarcinogenesis and might result in concomitant up-regulation of iNOS and COX II in the surrounding liver in favour of tumor promotion.

**7**
**IDENTIFICATION OF A METHYLATION-SPECIFIC MARKER ASSOCIATED WITH THE RENAL CELL CARCINOMA**

Barbara Seliger[1,2], Diana Handke[2], Rolf Damann[3], Rudolf Lichtenfels[1,2], Derek Atkins[1], Jürgen Bukur[1,2], Thomas Halder[4], Michael Kersten[4], Angelika Ackermann[1], Walburgis Brenner[5], Roland Kellner[6] and Friedrich Lottspeich[7]

[1]Johann-Gutenberg-Universität Mainz, III. Medizinische Klinik, 55131 Mainz;
[2]Martin-Luther-Universität Halle-Wittenberg, Institut für Medizinische Immunologie, 06112 Halle;
[3]Institute of Human Genetics, 06112 Halle;
[4]TopLab, 82152 Martinsried;
[5]Johann-Gutenberg-Universität Mainz, Klinik für Urologie, 55131 Mainz;
[6]Merck KgGA, 64271 Darmstadt; [7]Max-Planck-Institut für Biochemie, 82152 Martinsried, Germany

Renal cell carcinoma (RCC) represents the most common neoplasia of the kidney in Western countries. However, the molecular mechanisms leading to the initiation and the progression of this disease are still not well understood. Thus, novel innovative molecular markers for the diagnosis, prognosis as well as for the monitoring of this disease during therapy are urgently needed. Using proteome-based strategies for the identification of RCC-associated markers, both classic proteome analysis and PROTEOMEX, a combination of proteomics and serology, were employed for evaluation of the protein expression profile of normal kidney epithelium and RCC lesions. These approaches allowed the identification of a series of overexpressed and down-regulated proteins in RCC lesions in comparison to

normal kidney epithelium. Some of these genes have been implicated in tumorigenesis, but never linked to the initiation and/or progression of RCC. In particular, we found a marker which is also differentially expressed in other human epithelial tumors. Since tumor-associated antigens/markers are often repressed due to epigenetic mechanisms, the methylation status of this target was analyzed. A high frequency of total or partial methylation was detected in primary tumor lesions when compared to normal kidney epithelium and metastases. Currently, the functional significance of this proteins is being investigated in RCC.

## 8
## THE BRAIN PROTEOMICS OF NCL DISORDERS

Annina Lyly[1], Maria Lindfors[2], Carina von Schantz[1], Jaana Tyynelä[2] and Anu Jalanko[1]

[1]National Public Health Institute, Department of Molecular Medicine, Biomedicum Helsinki, P.O.Box 104, FIN-00251 Helsinki;
[2]Institute of Biomedicine/Biochemistry and Neuroscience Research Program, P.O.Box 63, FIN-00041 Helsinki, Finland; e-mail: annina.lyly@helsinki.fi

The neuronal ceroid lipofuscinoses (NCLs) comprise a group of disorders which deteriorates children's central nervous system. The incidence of these hereditary disorders is 1:12,500 worldwide. A common feature of all NCLs is progressive neurodegeneration, which leads to loss of vision, psychomotoric retardation and early death. In addition, lysosomal storage material accumulates in the tissues of the patients. Mutations in at least eight genes (*CLN1-CLN8*) underlie these diseases.

Infantile NCL (INCL) is the most severe form and is prevalent in the Finnish population. The symptoms start around the age of one year and, after rapid progression of the disease, the EEG of the patient is flat at three years of age. The defective gene, *CLN1*, encodes a soluble enzyme, palmitoyl protein thioesterase 1 (PPT1). PPT1 is lysosomal in non-neuronal cells, but in neurons it has been shown to localize to presynaptic axonal regions. Mutations in the *CLN5* gene cause a slower type of disease, where the first symptoms occur between 5 and 7 years of age. Similar to PPT1, CLN5 protein is also lysosomal in non-neuronal cells, but its function is yet unknown.

This study concentrates on the brain proteomics of two knock-out mouse models, *Cln1* and *Cln5*. We compared the protein expression levels of affected and control mice and searched for common characteristics between the two models. The methodology included 2D-electrophoresis of soluble cerebral sample combined with the identification of differently expressed proteins by mass spectrometry. We compared these data with the existing mRNA expression data from these mice to create a broader view of the metabolic pathways these diseases affect.

We have also created 2D-protein databanks of the analyzed mice brains and used different statistical analysis methods, such as PCA analysis, to compare the protein expression patterns between the different mouse models and controls. To verify the data from the proteomic analysis, we used biochemical methods, for example immuno-histochemistry and Western blotting.

## 9
## CATALOGUING THE SECRETOME OF SW620 HUMAN COLON CARCINOMA CELLS *VIA* 2D-PAGE AND MALDI-TOF/TOF-MS

Hanna Diehl[1,2], K. Stühler[1], Martin W. Volmer[1,2], Susanne Klein-Scory[2], Anna Schöneck[2], Wolff Schmiegel[2], Helmut E. Meyer[1] and I. Schwarte-Waldhoff[2]

[1]Medical Proteom-Center and
[2]Department of Internal Medicine (IMBL), Knappschaftskrankenhaus, Ruhr-University Bochum, Bochum, Germany

One major goal in cancer research is the identification of novel protein biomarkers to improve cancer diagnosis and monitor the course of disease. The direct detection of such markers in human blood or serum is fairly difficult due to the relatively low abundance of those proteins in comparison to the amount of "standard" serum proteins. Arguing that proteins released by tumour cells *in vitro* to a certain extent may reflect protein release *in vivo,* we started to analyse the tumour cell "secretome" as an alternative source for cancer markers. The word "secretome" may be tentatively defined as all proteins released by the cells into the conditioned media. As knowledge concerning the composition of such "secretomes" is limited, the aim of this study was to establish a catalogue of secretome proteins derived from SW620 human colon cancer cells, using modern proteomic techniques.

A previously established protocol was used for the production of the supernatant under serum-free conditions, an absolute prerequisite for the analysis of secretome proteins. Subsequently the proteins were harvested, concentrated and resolved by 2D-PAGE prior to visualizing the resulting spot pattern by silver-staining. The isolated proteins were in-gel digested with trypsin and the extracted peptides analysed by matrix-assisted laser desorption/ ionization time-of-flight mass spectrometry (MALDI-TOF/TOF-MS). Protein identification was achieved by submitting the mass lists of the measured peptide mass fingerprints (PMF) and peptide fragment fingerprints (PFF) to a database search applying Profound and Sequest search

algorithms. Using this approach, more than 60% (229 out of 355) of the proteins were identified. A first attempt was made to address the putative origin of the identified secretome proteins. This was performed by functional profiling of the results by database and published literature searches and through the analysis of leader sequences using the SigCleave program.

The SW620 secretome catalogue can thus be used as a basis to discover novel serum biomarkers. It also serves for comparison with secretomes derived from other (tumor) cells.

**10**

**PROTEOME ANALYSIS OF GASTRIC CANCER: IDENTIFICATION OF PROTEIN KEYS FOR THE DISCRIMINATION OF TUMOR AND NORMAL TISSUES**

Jae Won Kim[1], Euy Hoon Suh[2] and Chang-Won Lee[1]

[1]Laboratory of Proteome Analysis and Molecular Physiology, Department of Microbiology, Research Institute of Life Sciences and
[2]Department of Statistics and Data Analysis, College of Natural Sciences, Gyeongsang National University, 660-701 Jinju, South Korea

Gastric cancer remains a great challenge for clinicians and scientists. It is one of the most frequent cancers worldwide, and it is the second most common cause of cancer-related deaths. Patients with gastric cancers have a poor prognosis and low survival rates. Recently, we established a proteome database of gastric cancer tissues to analyze and mine the database and extract useful information on the proteins related with diverse aspects of gastric cancer. A database of the images of two-dimensional gel electrophoresis of tumor tissues and paired normal tissues of one hundred and fifty-two gastric cancer patients has been established. Paired tissue samples were obtained from 152 gastric cancer patients in the years 2001-2002 in Gyeongsang National University Hospital and Inha University Hospital, South Korea. Proteome analysis of gastric tissues was done by two-dimensional gel electrophoresis employing immobilized pH gradient (pH 4-7) isoelectricfocusing in the first dimension and 7.5-17.5% gradient SDS-PAGE in the second dimension. Protein visualization was done by staining with silver nitrate. The images were transformed into digital data by scanning and analyzed by PDQuest software for the automatic spot detection. On average, about one thousand protein spots were detected on the gels.

In this study, we investigated the proteome database and tried to identify protein keys useful for the discrimination of tumor tissue from normal tissue in gastric cancer. The data were analyzed by several statistical methods for the extraction of protein keys that are most determinative for the differentiation of tumor from normal gastric tissues. The protein keys thus extracted are presented and the different statistical methods employed for the analysis are compared and discussed.

**11**

**LC-MALDI-MS/MS ON A DISPOSABLE MALDI TARGET PLATFORM**

Martin Schuerenberg[1], Stephanie Hahner[1], Christian Ziegmann[2], Sven Brand[1], Wolfgang Jabs[1], Lars Vorwerg[1], Arnd Ingendoh[1] and D. Suckau[1]

[1]Bruker Daltonics, Fahrenheitstr. 4, D-28359 Bremen;
[2]Eppendorf AG, Hamburg, Germany

LC-MALDI on TOF/TOF instruments is one of the current breakthrough proteomics technologies. Since the chromatographic separation and the MS and MS/MS analyses are decoupled, the MS/MS parent ion selection process is freed from the time constraints in that decision making, that the LC-ESI-MS/MS process suffers from. After intelligent decision making has been completed, MS/MS acquisitions can be performed in a quite targeted fashion. As the chromatogram is "developed" as a permanent profile on a MALDI target, fractions can even be analyzed after further data acquisition-analysis cycles.

We applied the recently developed "Prespotted Anchor Chip" technology (PAC) to LC-MALDI analysis. PACs are hydrophobic plastic MALDI targets that contain small patches of hydrophilic spots covered by matrix thin layers. As these targets are designed for single use, memory effects are eliminated and also target archiving is possible. It was demonstrate that entire gradients developed onto PAC targets could be archived at room temperature for extended periods of time (>1 Month) with only minimal loss of information.

The basic target technology and its application to LC-MALDI analysis of proteins to achieve high sequence coverage and to detect posttranslational modifications are discussed.

**12**

**PRO-ANGIOGENIC AND PRO-INVASIVE EFFECT OF IRRADIATION AND COMBINED TEMOZOLOMIDE-RADIATION TREATMENT ON SURVIVING FRACTIONS IN HUMAN MALIGNANT GLIOMA CELLS**

Olga Golubnitschaja

Department of Radiology, Division of Molecular/Experimental Radiology, University of Bonn, Sigmund-Freud-Str. 25, D-53 105 Bonn, Germany

*Objectives*: Gliomas are the most common brain tumours, and grade IV glioblastomas are almost universally fatal. The current chemo-therapeutic treatment of glioma patients has minor success and little is known about the high resistance of gliomas towards existing therapies. This study investigated the effect of both temozolomide treatment and/or irradiation on the differential protein expression pa-tterns of surviving human malignant glioma cells. *Materials and Methods*: Morphological alterations under irradiation, chemo- and combined chemo/radio-treatment conditions were *in vitro* characterised on the genetically well-defined p53-wild-type U87Mg glioma cells. Cell proliferation and DNA damage were analysed using the "Comet Assay". Differential protein expression pa-tterns were investigated by 2D-PAGE and, in the selected spots, the protein products were identified by MALDI-TOF analysis. The distinguished protein expression rates were additionally quantified using Western blot analysis. Extracellular gelatinase activities for both MMP-2 and MMP-9 were determined by zymogramms in cell cultures for single incubation conditions. *Results and Conclusion*: Despite induction of extensive DNA damage, the clinically relevant treatments only partially suppressed the proliferation events in malignant glioma cells. Modulated expression patterns of MMPs/TIMPs, vimentin, lysosomal hydrolase C, and RhoA GTPase indicated a proangiogenic effect and an increasing grade of malignancy in irradiated and temozolomide-treated cell fractions. This conclusion correlates well with the extremely aggressive tumour phenotypes observed clinically at recidives of treated glioblastomas. Qualitatively, the new approaches should take the place of the current ones for the therapy of glioblastoma patients.

## 13
## PREDICTIVE DIAGNOSTIC MARKERS TO TARGET EXISTING CANCER THERAPIES

Christoph Petry

Bayer HealthCare AG, Diagnostics Research Germany, 51368 Leverkusen, Germany

Over recent years, medical progress has brought about significant advances the management of most cancers. Large scale clinical trials and ever refined therapeutic guidelines now allow for a stratification of patients towards a much more personalized treatment. Still, therapy responses to a given regimen are still highly variable, with the optimal response to an anti-cancer therapy often staying an illusive goal.

The most recent process in analytical tools, especially high throughput genotyping, now allows for recording the molecular fingerprint of a tumor, given a reasonable time and effort. The fingerprint can be correlated with therapy outcome, permitting the identification of biomarkers that predict therapy response. Ultimately, predictive markers could remove the need for iterative therapy optimization – a process with severe consequences for the mortality and morbidity associated with cancer. Especially for new targeted therapies, this approach has raised high hopes. However, recent experiences with new drugs show that the new concept continues to meet numerous obstacles.

Nevertheless, existing therapies appear to be a unique background for developing individualized treatment. Significant cohorts of patients are being treated with a well-controlled therapy regimen and clinical response is routinely monitored. Accordingly, the identification of predictive biomarkers for these therapies is an achievable goal with a potentially significant near- to mid-term impact on cancer care. Bayer HealthCare Diagnostics has embarked on a series of research projects to identify predictive diagnostic markers for breast cancer. Employing multiplex marker algorithms, Bayer HealthCare aims to provide products to help oncologists and pathologists make therapy decisions otherwise difficult using currently established diagnostic procedure.

## 14
## IDENTIFICATION OF THE FIRST CANDIDATE MOLECULAR MARKERS OF PANCREATIC CANCER PROGRESSION USING DIGE SATURATION LABELLING ANALYSING 1000 MICRODISSECTED CELLS

B. Sitek[1], J. Lüttges[2], K. Marcus[1], G. Klöppel[2], W. Schmiegel[3,4], H.E. Meyer[1], S.A. Hahn[3] and K. Stühler[1]

[1]Medical Proteom-Center, [3]Department of Internal Medicine, Knappschaftskrankenhaus and
[4]Department of Internal Medicine, Bergmannsheil, Ruhr-University Bochum;
[2]Department of Pathology, Christian Albrechts University, Kiel, Germany

In order to identify new molecular markers for pancreatic intraepithelial neoplasia (PanIN), the precursor lesions of pancreatic ductal adenocarcinoma, we established a proteomics approach analysing microdissected PanIN cells. Due to the limited amount of proteins available from microdissection, we developed a procedure including fluorescence dye saturation labelling in combination with high resolution two-dimensional gel electrophoresis (2-DE). With this procedure, we are able to analyse proteins extracted from 1,000 microdissected cells with a high resolution of up to 2,500 protein spots. Using protein lysates from pancreatic carcinoma tissue as a reference proteome, we successfully identified the proteins. Thus, approximately 2,200 protein spots (92%) of the microdissected sample proteome could be matched to the reference proteome for protein identification using MALDI-TOF-MS and nanoLC-ESIMS/MS after in-gel digestion. The first proteome analyses of PanIN-2 grades and

normal pancreatic ductal each revealed eight differentially-expressed proteins. The differential expression of the three actin filament-associated proteins transgelin, vimentin and MRLC3, as well as actin itself, shows a relevant role of the actin cytoskeleton during pancreatic tumour progression. Additionally, two members of the annexin family (annexin A2 and annexin A4) implicate a functional contribution of exocytotic and endocytotic pathways at that stage.

**15**
**BIOSYNTHETIC SPIDER SILK – THE ROUTE TO NOVEL ORTHOPAEDIC IMPLANTS**

Wilhelm Engström, Stefan Grip, Anna Rising, Helena Nimmervoll, Erica Storckenfeldt Margareta Stark, Göran Hjelm, Jan Johansson, Kefei Hu, Bror Morein, Stig Lindequist, Jan Skoglund and Armando Arias

Swedish University of Agricultural Sciences, Uppsala, Sweden

The need for improved implants in orthopaedic surgery has forced researchers to look for unconventional approaches for finding new biomaterials. Spiders spin up to seven different type of silk and each type possesses different mechanical properties. Reports on nucleotide sequences of cloned spider silk genes have gained importance as the mechanical properties of silk fibers have been revealed. It has, for example, been possible to link molecular data, often translated into amino acid sequences and predict three-dimensional structural motifs to known mechanical properties. Spider silks are principally composed of proteins with a predominance of glycine, alanine and serine. Silk proteins are able to undergo irreversible transformation from soluble protein to insoluble fibre. We have performed molecular studies on dragline silk proteins from the South African fishing spider *Euprosthenops senegalensis*. We report here sequence data linked to results from biomechanical studies, as well as the expression of spider silk cDNA clones in eukaryote cells.

**16**
**PROTEOMIC ANALYSIS OF GROWTH FACTOR SIGNALING PATHWAYS**

A. Saridaki, M. Samiotaki and G. Panayotou

Protein Chemistry Laboratory, Biomedical Sciences Research Center "Alexander Fleming", Vari, Greece

Growth factors interact with specific cell-surface receptors, which have intrinsic protein tyrosine-kinase activity and elicit cascades of intracellular signaling events, culminating in cellular growth. Proteomic technology can be applied to the detection of qualitative and quantitative changes of proteins that occur upon activation of these signaling pathways. We have established several systems that are amenable to proteomic analysis, including the comparison of growth factor-stimulated and resting cells, the overexpression of oncogenic receptors and the use of transgenic mice in which specific signaling components have been altered. The issues addressed include activation and inhibition of cell growth, cellular differentiation and oncogenic transformation. Using two-dimensional electrophoresis followed by mass spectrometric analysis, we have identified a number of proteins, several of which had not previously been implicated in the pathways activated by tyrosine kinase receptors. Subcellular fractionation has enabled us to identify changes due to intracellular protein translocation, in addition to post-translational modification. Finally, two distinct affinity purification strategies were employed for the isolation of protein complexes involving signaling molecules, followed by the identification of the interacting components using mass spectrometry.

**17**
**STIMULUS-INDUCED PLASTICITY OF THE OLFACTORY RECEPTOR NEURON**

Jon Barbour[1], Nadine Stoepel[3], Barbara Sitek[3], Kai Stühler[3], Dirk Wolters[2], Helmut E. Meyer[3], Bettina Warscheid[3], Hanns Hatt[1] and Eva Neuhaus[1]

[1]Lehrstuhl für Zellphysiologie, [2]Lehrstuhl für Analytische Chemie and [3]Medizinischen Proteom-Center, Ruhr-Universität Bochum, Germany

Olfactory receptors (OR) are G-protein-coupled membrane receptors that encode the largest vertebrate multigene family (~1,000 ORs in the mouse and rat, ~500-750 in humans); they are expressed individually in the sensory neurons of the nose and have also been identified in human testis and sperm. OR signal transduction is facilitated *via* adenylyl cyclase up-regulation of cAMP culminating in the opening of cyclic nucleotide-gated cation channels at the cell surface, which elicits a graded receptor potential.

We examined short-term and long-term olfactory receptor plasticity using established biochemical techniques in conjunction with a relatively novel proteomic strategy. Firstly, the question of short-term plasticity (*e.g.* receptor desensitization) was addressed by identifying novel OR receptor-protein-protein interactions using tagged fusion peptides. Tagged peptides from the intracellular loop 3 (IC3) and carboxyl termini of various mouse OR were used as bait to pull-out binding partners from mouse olfactory epithelium (OE). Interaction partners were identified using LC-MS/MS. In order to determine how OR mediate odour perception and how they influence long-term neuronal responses, a differential proteomic strategy was employed. Test mice were exposed to odorants and the OE was compared to the OE

from control mice using Fluorescence 2-D Difference Gel Electrophoresis (DIGE). Statistically significant differences in protein expression and peptide identification were found using DIGE and MALDI- TOF/LC-MS/MS, respectively. Both proteomic strategies afford a powerful means whereby novel protein-protein interactions can be elucidated and thereby provide greater insight into olfactory receptor plasticity.

**18**

**PROTEOMIC ANALYSIS OF ECTOPIC C-MYC OVEREXPRESSION *IN VIVO***

Vasiliki Ifandi[1], Susan E. Slade[2], Konstantinos Thalassinos[2], Sylvie. S. Abouna[1], Stella A. Pelengaris[1] and Michael Khan[1]

[1]Molecular Medicine Research Centre and [2]Biological Mass Spectrometry and Proteomics Group, Department of Biological Sciences, University of Warwick, Coventry, U.K.

In response to cellular signals, c-Myc (Myc) can regulate an extensive array of biological activities. It has been shown to induce cell cycle progression and proliferation, cell growth, potentiate apoptosis, block differentiation, drive transformation, activate genomic instability and stimulate angiogenesis. Myc regulates the expression of a large number of genes that, in turn, regulate multiple downstream events involved in all of these processes; however, at what point and to what extent these 'Myc-regulated' genes and their protein products diverge during the progression of various Myc-induced phenotypes *in vivo* is poorly understood. Given the ubiquitous role of Myc in human cancers as well as normal physiology, understanding how Myc can regulate such diverse and sometimes even opposing phenotypes (such as replication and apoptosis) is of central importance. In order to address these questions, we have developed a mouse model, which allows targeted expression of an inducible form of the Myc protein in various tissues *in vivo*. In one such model, expression of a regulatable chimaeric protein, MycER™, was targeted to suprabasal keratinocytes of the skin epidermis using the involucrin promoter. Continued activation of Myc (by topical administration of the activating ligand 4-hydroxytamoxifen) for up to three days results in relentless keratinocyte replication, accompanied by clear changes in protein expression as compared to wild-type tissue. Using a proteomic analysis approach incorporating two-dimensional gel electrophoresis, matrix-assisted laser desorption/ ionisation time of flight mass spectrometry and a variety of online published databases, these specific changes in protein expression were characterised. Proteins of diverse function were identified, including those involved in stress response, cell cycle, metabolism and maintenance of structural integrity.

**19**

**ART-LOVING BUGS: THE RESURRECTION OF SPINELLO ARETINO FROM PISA'S CEMETERY**

Paolo Antonioli[1], Giacomo Zapparoli[1], Pamela Abbruscato[2], Claudia Sorlini[2], Giancarlo Ranalli[3] and Pier Giorgio Righetti[1]

[1]Department of Industrial and Agricultural Biotechnologies, University of Verona, Strada Le Grazie 15, Verona 37134; [2]Department of Food Science and Microbiology, University of Milano, Via Celoria 2, Milano 20133; [3]Department of Food, Agrarian and Environment Science, University of Molise, Campobasso 86100, Italy

Between 1391 and 1392, Spinello Aretino painted a cycle of frescoes in Pisa's cemetery on the theme "Storie dei Santi Martiri Efisio e Potito", highly prized by contemporaries and by Vasari. Twenty years ago, one of these frescoes, entitled "Conversione di S. Efisio e Battaglia", due to discolouring and severe damage caused by humidity and atmospheric pollution, was removed from the walls by the "tear-off" technique, consisting of covering the surface with a strong cloth bound to it with generous layers of formaldehyde-treated glue. Subsequently, this large fresco (3.50 x 7.80 m) was abandoned in a storehouse for more than 20 years. When the careless curators attempted to remove the cloth, they found that the glue resisted any attempt at digestion, even when treated with the most aggressive proteolytic enzymes available on the market. Probably, during the long storage, the glue had been slowly cross-linked by the formaldehyde to the point of forming an intricate mass of untreatable proteinaceous material. Thus, although poor Spinello had died, in his own time, presumably as a *bona fide* Christian, his painting had been hidden for the rest of its life. When the fresco was recently treated by us with a suspension of viable *Pseudomonas stutzeri* cells, these bacteria were able to fully digest the hardened glue and restore Spinello's glorious painting to life. It is here shown how proteomics helped us to solve the riddle of how these bacteria acted on the obscured fresco.

**20**

**PREVIOUSLY UNDETECTED PROTEINS IN *ESCHERICHIA COLI* REVEALED BY ISOELECTRIC PREFRACTIONATION AND TWO-DIMENSIONAL GEL ELECTROPHORESIS**

Gary Smejkal

Proteome Systems, Woburn, MA, U.S.A.

Of the 4,377 genes transcribed in *Escherichia coli*, 4,290 are translated into proteins. This number of proteins is

increased to an undetermined number following post-translational modifications. However, fewer than 3,000 of these proteins have been isolated by two-dimensional gel electrophoresis (2DGE). This is due to the broad concentration range over which proteins are expressed, spanning at least six orders of magnitude, such that very low abundance proteins are not detected against the "background" of medium to high abundance proteins. Preparative solution phase isoelectric focusing (IEF) of *E. coli* lysate in the multicompartment electrolyzer (MCE) is capable of generating fractions of 0.5 and 1.0 pI intervals across the pH range 3-11. Concentration and individual analyses of these fractions on corresponding narrow range immobilized pH gradients (IPGs) detected four times as many proteins as were detected in the unfractionated bacterial lysate.

**21**

**"MULTIPLEXED PROTEOMICS" – NEW TECHNOLOGIES FOR ENHANCED PROTEIN IDENTIFICATION AND 'LABEL-FREE' SIMULTANEOUS QUALITATIVE/QUANTITATIVE PROTEIN PROFILING AND BIOMARKER DISCOVERY**

Alistair Wallace and Matt Kennedy

Waters Corporation, U.K.

Two novel systems that overcome key challenges in qualitative and quantitative proteomics are introduced.

Talk 1. Post-source decay (PSD) analysis of PSD fragment ions by MALDI-Tof is commonly used to provide extra specificity to increase the confidence of protein identification. Here, we describe a new MALDI MS/MS technique called "parallel PSD", which gives MS/MS information from all precursor ions ***in parallel*** (simultaneously), not in series, therefore enabling all the MS/MS data from one sample well to be acquired in just a few minutes. Results from this parallel PSD technique will be compared with MALDI/MS/MS analysis on a Q-Tof mass spectrometer.

Talk 2. A new LC/MS system for simultaneous qualitative and quantitative profiling of proteins from complex mixtures will be presented. The Waters® Protein Expression System employs an MS analyzer to acquire peptide and fragment mass data from multiple peptides ***in parallel***, to deliver "Global Expression Datasets" with significantly enhanced sampling efficiencies for increased protein sequence coverage. Furthermore, in quantitative proteomics, most contemporary approaches (*e.g.* ICAT) necessitate the chemical modification and analysis of peptides in a sample. The LC/MS platform presented does not require peptide derivatization, and quantitatively profiles the vast majority of the peptides within a sample to provide outstanding quantification statistics, in addition to enhanced protein identification in the same LC/MS run.

**22**

**MEMBRANE PROTEOMICS – CURRENT TOOLS AND METHODIC DEVELOPMENTS**

Isam Rais, Bjoern Meyer and Michael Karas

Institute of Pharmaceutical Chemistry, JW Goethe-University of Frankfurt,
Marie-Curie-Str. 9-11, D-60439 Frankfurt, Germany

Supercomplexes of the bovine mitochondrial respiratory chain, which can be isolated by Blue-Native (BN) electrophoresis and contain some 70 proteins, form an ideal system to test current technology and evaluate new methods for membrane proteomics, since they contain the different classes of membrane proteins, including some extremely hydrophobic (mitochondrially encoded) proteins, which are dominated by transmembrane helices. While the investigation of sliced 1D-SDS-PAGE gels by HPLC-Electrospray(ESI)-MS/MS analyses proved to be a generally applicable tool in membrane proteomics, a reliable gel electrophoretic separation and visualization of membrane proteins is still extremely valuable for accessing the composition and the constituents of the proteome investigated. We, therefore, developed a new, two-dimensional SDS-PAGE (doubleSDS) technique applying differently concentrated SDS gels, namely 10% in the first and 16% in the second dimension (or *vice versa*). Thereby, the extremely hydrophobic components were also clearly focused into spots with the extremely hydrophobic species located out of the gel diagonal. MALDI peptide mass fingerprinting for most of the constituents was easily carried out, although some of the extremely hydrophobic and low-molecular-mass integral membrane proteins needed some extra treatment, both by using alternative enzymes and improved extraction protocols. These approaches were compared to two alternative strategies, appyling tryptic digestion directly to BN bands and first using HPLC-ESI MS/MS and second using MALDI-MS/MS, after a simple reversed-phase separation and extraction.

**23**

**THE COMPARISON OF MICRODISSECTED TISSUES AND BIOFLUIDS BY MASS SPECTROMETRY**

Theo Luider

University of Rotterdam, The Nederlands

Microdissection permits the isolation of single cells or single populations of cells from thin tissue sections (5-10 μm). The whole microdissection procedure consists of two stages; laser microbeam microdissection followed by laser pressure

catapulting (LPC). Once the desired region is dissected from the surrounding tissue by a nitrogen laser beam, another strong laser is used to catapult the microdissected material directly into the tube cap used as a collection vessel. In combination with Mass Spectrometry (MS), this technique allows the acquirement of protein signatures from 100-200 cells from a frozen tissue sample. Protein or peptide identification has been performed using offline nano Liquid Chromatography Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (nano-LC MALDI-TOF/TOF) mass spectrometry. Comparison of peptide signatures of microdissected cells resulted in significantly different peptides among healthy and diseased tissue sections (diseases studied were preeclampsia and breast carcinoma). Development in this innovative research area allows the opportunity to determine specific peptides and protein pathways localized in certain cell types in relation to the disease studied. We now show that this method is potentially suitable for finding specific peptides and for the identification of proteins that are related to pathogenesis of trophoblast pregnancy diseases, such as preeclampsia (1) and the finding of differentially expressed proteins in breast carcinoma cells and their surrounding stroma cells. As a result, large data sets of peptides and proteins have been compared. Therefore, different statistic software programs have been tested to focus on disease-related peptides/proteins and to exclude proteins that relate to chance, patient variation and technical variations.

In a similar approach, cerebrospinal fluid (CSF) samples of breast carcinoma patients that had leptomeningeal metastasis were compared with patients with breast carcinoma without leptomeningeal metastasis and a control group of patients without cancer. Characteristic peptide profiles were identified using specific software tools. A qualifier was built and the differentially expressed peptides were identified by nano Liquid Chromatography Fourier Transform Ion Cyclotron Resonance (nano-LC FT-ICR) mass spectrometry.

1 De Groot CJM, Steegers-Theunissen RP, Güzel C, Steegers EAP and Luider TM: Peptide patterns of laser-dissected trophoblasts analysed by MALDI-TOF mass spectrometry. Proteomics *5*: 597-607, 2005.

**24**

## TOP-DOWN PROTEOMICS: ANALYSIS OF INTACT PROTEINS WITH LTQ FT

Michaela Scigelova[1], Tori Richmond[2], Jennifer Sutton[2] and Leo Bonilla[2]

[1]Thermo Electron, Hemel Hempstead, U.K.;
[2]BRIMS Center, Thermo Electron, Cambridge, MA, U.S.A.

Protein identification based on the analysis of intact proteins has many advantages over methods relying on the analysis of peptides after enzymatic digestion of the original biomolecules. Measuring the protein in its intact form captures it in the context of its co- and post-translational modifications, and life-span alterations which might be otherwise missed by other approaches.

The presentation focuses on various aspects of intact protein analysis. The benefit of high mass accuracy and ultra-high resolution obtainable from FT ICR instrumentation is illustrated. An example of yeast proteome analysis demonstrates the use of ProSight PTM software for the identification of intact proteins by database searching. Direct coupling of nano-LC with the FT ICR system in real time is discussed in detail as used for the automated analysis of low molecular weight human plasma proteome.

**25**

## STRUCTURAL STUDIES OF NONCOVALENT INTERACTIONS BETWEEN β-AMYLOID PEPTIDE AND BIOACTIVE COMPOUNDS BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

Fotini N. Bazoti[1,2], Jonas Bergquist[3], Karin Markides[3] and Antonis Tsarbopoulos[1,2]

[1]Pharmacy Department, Laboratory of Instrumental Analysis, University of Patras, Patras 265 04;
[2]GAIA Research Center, Bioanalytical Department, The Goulandris Natural History Museum, 13 Levidou Street, Kifissia 14562, Greece;
[3]Department of Analytical Chemistry, Institite of Chemistry, Uppsala University, Box 531, SE-751 21, Uppsala, Sweden

Cellular functions are often triggered by weak noncovalent interactions between proteins and ligands. Interruption of these noncovalent interactions could lead to disruption of normal cellular processes, which could lead to diseases. Therefore, elucidation of the structures and formation mechanisms of these complexes could offer a better understanding of the disease processes, which in turn may aid in modern drug design.

The introduction of electrospray ionization (ESI) mass spectrometry (MS) has revolutionized the application of MS towards the study of high molecular weight proteins, either by themselves or bound to ligands. Monitoring of protein interactions in real time by ESI MS allows the definition of the stoichiometry and the topology of the interacting species. This is well illustrated in the analysis of the *ras*-GDP and *ras*-GTP complexes, which are important in cancer research. These observations have formed the basis for screening potential inhibitors that bind and inactivate the *ras* protein.

The formation of noncovalent complexes of amyloid beta peptide (Aβ) with bioactive compounds was also evaluated. The Aβ peptide is currently believed to play a central role in

the pathogenesis of Alzheimer's disease (AD), even though the cause of AD is unknown. In light of the suggested link between oxidative stress and AD, it is proposed that antioxidants, and even more endogenous antioxidants, such as Melatonin (M), may offer a therapeutic regime for protection against this disease. We have studied the formation of noncovalent complexes of Aβ with M and certain bioactive phytochemicals, derived from plants endemic in Mediterranean flora, by ESI MS. Moreover, the localization of the ligand binding sites was assessed by employing digestion protocols of the respective Aβ noncovalent complexes combined with FTICR-ESI MS analysis. The stoichiometry of the Aβ noncovalent complexes was found to vary with incubation time. FTICR-ESI MS analysis of the Aβ–M complex proteolytic fragments indicates that the [4-11] sequence and the [17-28] hydrophobic region of Aβ are responsible for the noncovalent interaction. These data may shed light onto the mechanisms of AD pathology, and facilitate the design of novel aggregation inhibitors, thus aiding the prevention or treatment of AD.

**26**

## RULE-BASED APPROACH TO INTERPRET MS/MS SPECTRA APPLICATION TO THE DISCOVERY OF INNER MEMBRANE MITOCHONDRIAL PROTEINS

Ioannis Xenarios

Serono Pharmaceutical Research Institute
Plan-les-Ouates, Switzerland

The identification and characterization of proteins are parts of most biotechnology projects. Several computational tools have been developed over the years to analyze mass spectrometry-derived MS/MS spectra, but one aspect that has not been addressed is the validation of spectra. Probabilistic frameworks have been introduced to 'infer' accuracies to the interpreted MS/MS spectra. We have developed a rule-based method that takes advantage of criteria used by an experienced mass spectrometrist. The algorithm was tested for the characterization of proteins contained within the inner membrane of the mitochondria. This approach allowed the discovery of several novel inner membrane proteins.

**27**

## INVESTIGATION OF CHANGES AT THE PROTEOMIC LEVEL: OPTIMISATION OF SAMPLE PREPARATION AND MASS SPECTROMETRY ENHANCEMENT

Klaus Herick[1], Jodi M.Zobrist[2], Jennifer R.Cresswell[2], Justin Wildsmith[2], Richard J.Mehigh[2], Matthew V.Caple[2], Kelly Foster[2] and Rafael Valdes-Camin[2]

[1]Sigma-Aldrich Chemie GmbH, Eschenstr. 5, 82024 Taufkirchen, Germany
[2]Sigma-Aldrich Corporation, St. Louis, MO, U.S.A.

While there is no universal solution for both simplifying the complexity of protein samples from cells, tissues or organs and addressing protein solubility issues, a systematic approach of optimising sample preparation methods for 2-D electrophoresis (2-DE) is an extremely powerful tool. Therefore a variety of detergents were tested to determine which resulted in the greatest number of well-resolved spots in 2-DE for Chinese hamster ovary cells (CHO-K1).

After optimising the method for total protein extraction, various samples were pulled at different time-points from a bioreactor and studied by 2-DE. Using expression analysis software for comparison of the gels, changes in protein expression were quantitated. Selected gel spots were identified using MALDI-MS after in-gel tryptic digestion of the proteins. MS enhancements by guanidination for MS are presented.

By making use of optimised extraction methods, such as the one demonstrated for CHO-K1, 2-DE coupled with enhancements for MALDI-MS analysis becomes a more effective tool for studying changes in protein expression.

In addition, alternative approaches for protein expression profiling are discussed, including MS-based quantitative proteomics using the relative method of $^{18}O$ incorporation upon tryptic digestion and absolute quantitation using AQUA peptides as internal standards.

Cockrill SL *et al*: Efficient micro-recovery and guanidination of peptides directly from MALDI target spots. BioTechniques *38*: 301-304, 2005.

Gerber SA *et al*: Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS PNAS *100(12)*: 6940-6945, 2003.

Lunardi J *et al*: Improvement of the solubilisation of proteins in two-dimensional electrophoresis with immobilised pH-gradient. Electrophoresis *18*: 307-316, 1997.

Rabilloud T *et al*: Structure-efficiency relationships of zwitterionic detergents as protein solubilisers in two-dimensional electrophoresis. Proteomics *3*: 111-121, 2003.

Rabilloud T *et al*: Evaluation of nonionic and zwitterionic detergents as membrane protein solubilisers in two-dimensional electrophoresis. Proteomics *3*: 249-253, 2003.

**28**

## MALDI MASS SPECTROMETRIC IMAGING: A NEW TOOL FOR BIOMEDICAL RESEARCH

Tatiana C. Rohner, Dieter Staab and Markus Stoeckli

Analytical and Imaging Sciences, Novartis Institutes for BioMedical Research, CH-4002 Basel, Switzerland

In biomedical research, the discovery of new drugs demands analytical techniques with high sensitivity together with increased throughput. Localizing or following changes in organisms at the molecular level by imaging component

distributions of specific tissues can help to unravel biochemical pathways and to identify targets or biomarkers.

Most of the imaging techniques, such as MRI and PET, however, need reporters to probe analytes of interest, and thus simultaneous detection of different biomolecules is cumbersome and sometimes not feasible. The high sensitivity of mass spectrometry over a wide range of compounds has been associated to an automated MALDI MS analysis to allow the acquisition of molecular images based on MS spectra. To perform MALDI mass spectrometric imaging (MSI), the UV pulsed laser of the MALDI source is used to scan a selected area of a biological tissue deposited on a MALDI plate, while acquiring mass spectra of the ablated ions at every image point, *i.e.* each laser shot position. Based on the database of acquired spectra, signals are selected and analyte-specific images can be generated.

MALDI MSI can be used to track biomarkers, such as peptides or proteins, but also to map drug/tissue interactions. Results based on the molecular scanner approach, which gives access to high mass range by combining tissue blotting and digestion in a one-step process, is also presented.

**29**

## THE NOVEL STS MALDI BIOCHIP PLATFORM FOR BIOMARKER DISCOVERY

Lina P. Aristoteli, Michael Mariani and Mark S. Baker

Australian Proteome Analysis Facility Ltd., Department of Chemistry and Biomolecular Sciences, Research Park Drive, Macquarie University, Sydney NSW 2109, Australia

MALDI mass spectrometry sensitivity to the attomole range of peptides is needed for identification of low abundance proteins and peptides in the proteome. In addition, there is a need to develop techniques that minimize ion suppression caused by contaminants and sample complexity. The Surface Tension Segmented (STS) MALDI platform is based upon chemically-defined virtual wells, where some surfaces are highly ordered self-assembled monolayers containing grafted chemistries. The biochip enables the addition of up to 50 μL of analyte to each well by virtue of concentric zones of varying "wetability", allowing sample concentration to a 0.6 mm central analysis zone. The STS biochip range comprises the STS-RP3™ biochip to selectively remove hydrophilic contaminants from samples, relevant for MALDI analysis of trypsin digests. High affinity STS biochips utilize varying surface chemistries such as immobilized metals, antibodies, and/or streptavidin, to reduce the complexity of samples through affinity selection of analytes from complex samples. Our preliminary results show the novel application of this technology to the sensitive detection of isoforms of purified protein and/or peptide biomarkers (*e.g.*, PSA, CEA, CA-125, *etc*). Further, the utility of the platform allows for both MS and MS/MS data to be obtained by TOF-TOF mass spectrometry in the attomole per micro litre range (amol/μL) using the Applied Biosystems 4700.

APAF acknowledges the assistance of the LCI team including Chris Belisle, Doug Greiner and John Walker III.

**30**

## TOOLS FOR SAMPLE PREPARATION AND FRACTIONATION IN 2D ELECTROPHORESIS

Anton Posch, Aran Paulus and Mary Grace Brubacher

Bio-Rad Laboratories, Hercules, CA, U.S.A.

Two-dimensional gel electrophoresis (2DE), with its unique capacity to resolve thousands of proteins in a single run, is a fundamental research tool for nearly all protein-related scientific projects. Over the last two decades, numerous researchers in academia and in industry have improved the technology to the point where a novice user is capable of achieving respectable gel separations on the first try. In addition, 2DE technology, as practiced today, has seen enormous gains in reproducibility, resolution and automation, all of which contribute to its widespread use. Nevertheless, 2DE is still a technically demanding method.

The quest to map and characterize each and every protein in a given cell type, tissue or organism has given 2DE an additional boost as the separation method of choice for many proteomics laboratories. However, the task list for a proteomics researcher is daunting: the number of proteins in a biological sample, although unknown at this time, is believed to be in the 100,000s, covering a concentration range of 7 or more orders of magnitude. In addition, the proteome is extremely dynamic, with protein expression depending on the cell state and further complicated by posttranslational modifications such as phosphorylation or glycolysation, to name just two possible changes to proteins in a functional biological system.

As more and more laboratories start up their own proteomic effort or revamp existing programs, they realize that meticulous attention to 2DE methodology is only one critical aspect when identifying differentially-expressed proteins or investigating a particular biological pathway.

The information content of 2DE is heavily influenced by a correct sample preparation strategy. Interestingly, not much attention was paid to this area during 2DE methodology development.

Here we provide a broad overview of the principles and recent developments in sample preparation tools prior to the first step of 2DE. Examples from three strategies for sample preparation, based on solution chemistry, chromatography and electrophoresis, are discussed in detail,

and also used to illustrate how these key areas can be applied to general-purpose sample cleanup and sample fractionation for enrichment of low abundance proteins.

**31**
**TODAY'S 2-D ELECTROPHORESIS TECHNOLOGY**

A. Görg, O. Drews, C. Lück, G. Reil and W. Weiss

Technical University of Munich, Department of Proteomics, D-85350 Freising-Weihenstephan, Germany

Two-dimensional gel electrophoresis (2DE), with immobilized pH gradients (IPGs) combined with protein identification by mass spectrometry (MS), is currently the work horse for proteomics. Mass spectrometry has developed into a highly sophisticated technology, whereas immobilized pH gradients have overcome the former limitations of carrier ampholyte-based 2DE (O'Farrell, 1975) with respect to reproducibility, handling, resolution, and separation of very acidic and/or basic proteins (NEPHGE). The development of IPGs between pH 2.5-12 has facilitated the analysis of very acidic and very alkaline proteins and the construction of the corresponding databases (http://www. wzw.tum.de/proteomik). Moreover, narrow-overlapping IPGs provide increased resolution (pI=0.001) and, in combination with prefractionation methods, the detection of low abundance proteins. Sample prefractionation with IEF in granulated gels (Görg *et al*: Proteomics *2*: 1652-1657, 2002) has been successfully used for the enrichment of low-abundance proteins. Moreover, it also improved the quality of MALDI- MS spectra, both in terms of sequence coverage (number of peptides) and signal to noise ratio.

In spite of alternative technologies that have emerged (MudPIT, stable isotope labelling, arrays), 2DE is currently the only technique that can be routinely applied for parallel quantitative expression profiling of large sets of complex protein mixtures. Furthermore, it delivers a map of intact proteins, which reflects changes in protein expression level, isoforms or post-translational modifications. Last but not least, today's 2DE technology with IPGs (Görg *et al*: Electrophoresis *21*: 1037-1053, 2000), in combination with prefractionation techniques, DIGE and mass spectrometry, has greatly improved the coverage of the total proteome of a cell.

**32**
**PROTEIN MICROARRAY TECHNOLOGY: TECHNOLOGIES, APPLICATIONS AND MARKET**

Thomas Joos

NMI Natural and Medical Sciences Institute at the University of Tuebingen, 72770 Reutlingen, Germany

Following the completion of the human genome sequencing project, well-established DNA microarrays and sophisticated bioinformatics platforms allow scientists to take a global view of biological systems. In today's proteome era, the time is ripe for protein microarrays to screen entire genomes for proteins that interact with particular factors, catalyze particular reactions, act as substrates for protein-modifying enzymes and/or as targets of autoimmune responses. Due to miniaturization, microarrays can analyze many parameters in parallel, requiring only minimal amounts of reagents and sample. Besides planar microarray-based systems, bead-based flow cytometry approaches are very well suited to the multiplex detection of target molecules, especially when only a small number of parameters have to be determined simultaneously. Sensitivity, reproducibility, robustness and automation have to be demonstrated before this technology will be suitable for high-throughput applications within the field of proteomic research and bio-marker discovery. The current state of miniaturized and multiplexed approaches for protein microarray technology is summarized. The results of different types of assays generated at the NMI are presented and discussed.

**33**
**SELECTIVE LABELLING OF PLASMA MEMBRANE PROTEINS OF JURKAT T – CELLS USING CY DYE DIGE FLUOR MINIMAL DYES**

Caroline May[1], Stefan Helling[1], Elmar Langenfeld[1], Thomas Schulenborg[1], Helmut E. Meyer[1], Stefan Müllner[2], Petra Lutter[2] and Katrin Marcus[1]

[1]Medical Proteom-Center, Ruhr - University Bochum, Bochum;
[2]Protagen AG, Dortmund, Germany

Cell surface-associated proteins and integral membrane proteins of the plasma membrane, *e.g.* receptors, transporters and ion channels, play an important role in cellular mechanisms. Receptors in particular act as mediators for processes like survival, differentiation, division and cell death. Therefore, these proteins are a bountiful source of potential drug targets and represent an interesting topic of research, for instance for the pharmaceutical industry.

The aim of our study was to selectively label plasma membrane proteins of Jurkat cells *in vivo* using fluorescence dyes, facilitating a direct "in-gel" differentiation and quantification of plasma membrane proteins, without the need of blotting steps or even identification of all proteins by LC - MS/MS.

In this work, we describe a protocol for the differential analysis of plasma membrane proteins using cell surface

labelling of intact cells, employing fluorescence dyes (Cy3 and Cy5, GE Healthcare Bio-Sciences) followed by Triton X-114 extraction and phase partitioning (1). Due to the hydrophobic nature of membrane proteins, they mainly passed into the detergent phase. Subsequently, the labelled proteins were separated by either 1-D or 2-D PAGE and labelled proteins were detected using a fluorescence scanner (Typhoon™, GE Healthcare Bio-Sciences). Protein identification was done by nanoLC - ESI - MS/MS. Preliminary results with fluorescence microscopy clearly demonstrate that labelling only occurs at the cell surface. Separation and detection of both the detergent and the soluble protein fractions also demonstrated specific labelling of plasma membrane proteins. The identification by nanoLC - ESI - MS/MS showed that the labelled proteins in the detergent fraction were membrane or membrane-associated proteins. After separation by 2-D- PAGE, some labelled spots in the detergent fraction and also in the aqueous fraction could be detected. The mass spectrometric analysis of the protein spots is still ongoing. In the future, another 2-D gel system (similar to BAC/SDS-PAGE) will be used for a more effective separation and description of the membrane protein fraction.

1 Bordier C: Phase separation of integral membrane proteins in Triton X-114 solution. J Biol Chem *256(4)*: 1604-1607, 1981.

**34**

## A PREFRACTIONATION STRATEGY FOR PROTEIN PROFILING OF *LEISHMANIA MAJOR* PROMASTIGOTES BY TWO-DIMENSIONAL ELECTROPHORESIS

Hediyeh Rasam, Sarvnaz Hashemi, Alireza Gholami, Maryam Zakeri and Behrouz Vaziri

Protein Chemistry Unit, Biotechnology Research Center, Pasteur Institute of Iran Azad, Iran

Using proteomics for identifying the stage-specific proteins in digenetic infectious organisms has great potential in defining new vaccine candidates and drug targets. *Leishmania major* is the causative agent of coetaneous leishmaniasis in Iran.The purpose of this study was to achieve an efficient sample preparation method for profiling the proteome of *Leishmania major* promastigotes by two-dimensional electrophoresis (2DE). *Leishmania major* (MRHO/IR/76/ER) promastigotes were collected at the stationary phase and lysed by a buffer containing 5M Urea, 2M thiourea, 2% SB3-10, 2% CHAPS and 0.2% ampholyte optimized for solublizing the cell proteins. A pre-fractionation step was utilized based on reverse phase- high performance liquid chromatography (RP-HPLC). Through a step gradient elution, the fractions were collected and applied to immobilized pH gradient gels for IEF. In this study, we established an efficient sample preparation method for producing highly resolved and reproducible 2DE gels. Moreover, our experiments showed that adding a pre-fractionation step, as another dimension to the separation of the whole cell lysate, gives a high resolving power for separation of the parasite proteins.

**35**

## SEARCHING FOR HYPOTHETICAL PROTEINS: A NECESSITY

Gert Lubec, Leila Afjehi-Sadat and Jae-Won Yang

Medical University of Vienna, Department of Pediatrics, Austria

A large portion of mammalian proteomes is represented by hypothetical proteins (HP), *i.e.* proteins predicted from nucleic acid sequences only and protein sequences with unknown function. Moreover, databases are far from being complete and errors are to be expected.

The legion of HPs is awaiting experiments to show their existence at the protein level; subsequent bioinformatic handling in order to assign proteins a tentative function is mandatory.

Two-dimensional gel electrophoresis, with subsequent mass spectrometrical identification of protein spots, is an appropriate tool to search for HP in the high-throughput mode. Spots are identified by MS or ideally by MS/MS measurements (MALDI-TOF,MALDI-TOF-TOF) and subsequent software, *e.g.* Mascot, Profound and MS-fit, to name a few. In many cases, proteins can thus be unambiguously identified and characterized; if this is not the case, *de novo* sequencing or Q-TOF analysis is warranted. If the protein is not identified, the protein sequence is sent to databases for BLAST searches to determine identities/ similarities or homologies to known proteins. If no significant identity to known structures is observed, the protein sequence is then examined for the presence of functional domains (databases Pfam, *etc.*) and subjected to searches for motifs and finally protein-protein databases are consulted.

We here provide information about hypothetical proteins in terms of protein chemical analysis, independent of antibody availability and specificity and bioinformatic handling, to contribute to the extension / completion of protein databases and include original work on HP in the brain to illustrate the processes of HP identification and functional assignment.

**36**

## COMPARATIVE GENOMICS OF THE GENUS *DROSOPHILA* AND OTHER INSECTS

William M. Gelbart

Department of Molecular and Celluar Biology, Harvard University, Cambridge, MA 02138, U.S.A.

Much basic genetic and genomic research focuses on the analysis of a relatively few model systems, including among them the fruit fly, *Drosophila melanogaster*, While the direct analysis of the *D. melanogaster* genome has been extremely productive, it is clear that additional information is needed to understand many aspects of its gene organization and evolution, and to make the most rigorous possible predictions of its proteome. Another extremely valuable source of additional information are the genome sequences of 9 other species of *Drosophila,* arranged at different evolutionary distances from *D. melanogaster*. While analysis of these genomes is in its initial phases, it is clear that already it is assisting in the understanding of the encoded proteome, noncoding sequence conservation and chromosome architecture. These topics are discussed in the context of the genus *Drosophila* and extended to consider the comparative genomics needs for other important insect targets, such as vectors of disease, agricultural pests and other experimental models.

**37**
**LIVING TURING MACHINES AND THE DELPHIC BOAT**

Antoine Danchin

Genetics of Bacterial Genomes, Institut Pasteur, Paris, France

Amongst the questions asked by the Pythia was this enigma: If we consider a boat made of planks, what is it that makes the boat a boat? As time passes, some of the planks begin to rot and have to be replaced. There comes a time when not one of the original planks is left. The boat still looks like the original one, but in material terms it has changed. Is it still the *same boat*? The owner would say yes, this is my boat. Yet none of the material it was originally built from is still there. If we were to analyze the components of the boat, we would not learn very much – a boat made from planks of oak is different from a boat made from planks of pine, but this is fairly incidental. What is important about the material of the planks, apart from their relative stability over time, is the fact that it allows them to be shaped, so that they relate to each other in a certain way. The boat is not the material it is made from, but something else, much more interesting, which organizes the material of the planks: the boat is *the relationship* between the planks. Similarly, the study of life should never be restricted to objects, but must look into their relationships. This is why a genome cannot and must not be regarded as simply a collection of genes. Because relationships between physical objects are at the core of life, the study of living organisms stems from symbolic abstraction. More often than not, the objects that create biological functions have no straightforward and "mechanical" coupling with them; they are only their mediator, their symbol. At the genome level, it is the relationship the genes have with each other, and with the signals that control them, that gives life to an organism. Within every cell, something is passed on from generation to generation, transmitting what common sense calls information, by a process which current models of heredity compare to the running of a program, in a sense that it is close to the way this word is used in computer sciences. In brief, one can consider cells as Turing machines. This creates a paradox, however, as noticed by John von Neumann in the early sixties: a Turing machine does not make Turing machines. How is this paradox resolved? Where is the operating system, and what does the machine read? This will be the topic of our discussion.

**38**
**FUNCTIONAL GENOMICS OF THE *ANOPHELES GAMBIAE* INNATE IMMUNE SYSTEM AND ITS INTERACTION WITH THE MALARIA PARASITE**

George K. Christophides

European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

*Anopheles* mosquitoes are obligatory vectors of the malaria parasite, *Plasmodium*, which causes an annual death toll of over two million people. Our work capitalises on the newly available genome sequence of *Anopheles gambiae*, the major vector of human malaria in Africa, and on postgenomic tools to understand the mosquito innate immune system and its interactions with the parasite. We have established that innate immunity is a major determinant of the mosquito's vectorial capacity, as it accounts for substantial parasite losses during invasion of the mosquito midgut. However, it also allows immune evasion leading to successful malaria transmission. We have elucidated key immune reactions, including parasite lysis and melanisation, at the molecular and cellular level, and have begun to dissect the regulatory networks underlying these immune reactions. This work contributes to the emerging comparative understanding of innate immune mechanisms and may open future possibilities for controlling malaria transmission.

**39**
**DATA MANAGEMENT AND BIOINFORMATICS SOLUTIONS FOR HIGH-THROUGHPUT PROTEOMICS**

Thodoros Topaloglou

Scientific Computing, Protana Inc., Canada

Large scale proteomics analysis of complex biological samples is rich in data management and bioinformatics

challenges. For example, biomarker discovery studies routinely involve hundreds of LC-MS acquisitions, producing spectral data that need to be transformed into quantitative expression profiles and compared to identify statistically significant and biologically relevant biomarker candidates. This presentation highlights challenges in computational proteomics and presents solutions that are part of a bioinformatics platform for an industrial mass spectrometry-based proteomics laboratory that focuses on biomarker discovery. The essential components of the platform are methods for accurate assignment of proteins to observed peptides, construction of peptide profiles across samples from raw mass spectra, statistical analysis for establishing significant expression profiles and biological interpretation of results. A distinguishing feature of the platform is the integration of these components through an underlying infrastructure for data and workflow management, which enables the linking of otherwise disparate information and processes from sample preparation, data acquisition, protein identification, spectral analysis, statistical analysis and biological interpretation. Specific examples are cited to demonstrate the throughput gains that can be delivered from the combination of comprehensive data management and an integration framework with computational tools for quantitative proteomics.

**40**
**DNA SEQUENCING AND GENERALIZED TRANSITION POLYNOMIALS**

Irasema Sarmiento

Cinvestav, Mexico

The interlace polynomial of a graph, introduced by Arratia, Bolloba's and Sorkin, has evolved from problems in DNA sequencing. They showed that the number of potential reconstructions in DNA sequencing may be modelled by Eulerian digraphs. Subsequently, Bouchet showed that the interlace polynomial is equivalent to a polynomial called Martin polynomial. The generalized transition polynomials of Ellis-Monaghan and Sarmiento explore the general structure of such polynomials, including certain knot and link invariants with their well-known connections to the behaviour of DNA strands. Our work is directly relevant to biomedical research in the area of DNA

**41**
**ENHANCING BIOMEDICAL RESEARCH WITH SYSTEMS LITERATURE ANALYSIS**

Christos Andronis

Biovista, Greece

Advances in experimental technology in the last decade have drastically changed the way biological knowledge is generated. Whole genome sequencing and large-scale genomic and proteomic experiments are providing researchers with an unprecedented amount of data that constantly needs to be interpreted, with the ultimate goal being the complete understanding of the inner workings of complex biological systems.

The generation of data at the experimental level is paralleled by an explosion of information published in scientific papers. NLM's Medline database now indexes more than 15 million abstracts and approximately 1,500 new abstracts are added every day to the database. By any standards, this is a huge amount of information, making it almost impossible for the unaided researcher to keep up with developments in his/her own field of research. A common issue with the use of some of the well-known resources (*e.g.* PubMed, Google) for bibliographic analysis is that, often, a lot of irrelevant information is returned, creating an additional workload for the researcher. It is, therefore, clear that a new model of bibliographic analysis is needed in order for researchers to take full advantage of the literature data.

Literature analysis can fit in the modern research pipeline, both by helping the researcher to exploit the flood of publications, but also by aiding the analysis of high-throughput assays. During recent years, there has been an increase of interest in using data mining techniques to extract information from the biomedical literature and present it to the researcher in a meaningful and assimilable manner. Systems Literature Analysis (SLA) is an emerging method in this area. SLA treats the literature as a system of interconnected research parameters (cell lines, genes, pathways, diseases, *etc*.) and uses text mining techniques to extract relationships and connections. In this respect, SLA can be considered to be the literature-driven version of systems biology.

The SLA concept, as well as ways to perform literature-based analysis of proteomic and genomic data, are presented, in the context of Biolab Experiment Assistant (BEA), an integrated literature analysis environment developed by Biovista.

**42**
**COMPUTATION OF LARGE PHYLOGENETIC TREES: ALGORITHMIC AND TECHNICAL SOLUTIONS**

Alexandros Stamatakis

Foundation for Research and Technology-Hellas, Greece

The computation of ever larger as well as more accurate phylogenetic trees, with the ultimate goal of computing the "tree of life", represents one of the grand challenges

in high performance computing (HPC) Bioinformatics. Statistical methods of phylogenetic analysis, such as maximum likelihood and Bayesian inference, have proved to be the most accurate models for evolutionary tree reconstruction.

Unfortunately, the size of trees which can be computed in a reasonable time is limited by the considerable computational cost induced by these methods. Two orthogonal research directions exist to overcome this challenging computational burden: firstly, the development of novel, faster and more accurate heuristic algorithms; and secondly, the application of high performance computing techniques, the deployment of supercomputers, and Grid-computing to provide the required computational power, mainly in terms of CPU hours.

The field has witnessed significant algorithmic advances over the last 2-3 years, which allow for inference of large phylogenetic trees containing 500-1,000 sequences on a single PC processor within a couple of hours using maximum likelihood. On the other hand, the main problem which high performance computing implementations of maximum likelihood analyses face is that technical development lags behind algorithmic development, *i.e.* programs are parallelized that do not represent the state-of-the-art algorithms any more.

Within this context, a brief overview of the computational challenges large-scale phylogenetic inference face, concerning both algorithmic as well as supercomputing aspects, is given.

Subsequently, the benefits of simultaneous algorithmic and technical development are outlined, using as an example the RAxML program (Randomized Axelerated Maximum Likelihood). The parallel version of RAxML has been used to compute the largest maximum likelihood tree to date (comprising 10,000 organisms) on a Linux PC cluster.

Finally, recent algorithmic developments, such as the application of fast simulated annealing techniques as well as divide-and-conquer strategies, are discussed. An overview of possible future HPC implementations of those novel algorithms is provided including Grid-based solutions, implementations for hybrid supercomputer architectures, and exploitation of vector-like peripheral processors *e.g.* Graphics Processing Units (GPUs).

**43**

## PROTEOMICS AND CLINICAL CARDIOLOGY: ACHIEVEMENTS AND PROSPECTIVES

P. Christopoulos[1], S. Kourtis[2], A. Mouzaki[1] and J. Christodoulou[2]

[1]Laboratory of Immunology, University of Patras
[2]2nd Cardiology Department, "St. Andrew" General Hospital of Patras, Greece

*Background*: Heart diseases resulting in heart failure are among the leading causes of morbidity and mortality and can result from either systemic disease (*e.g.* hypertensive heart disease, ischaemic heart disease [IHD]) or specific heart muscle disease (*e.g.* dilated cardiomyopathy [DCM]). The causes of cardiac dysfunction in most heart diseases are still largely unknown, but are likely to result from underlying alterations in gene and protein expressions. Proteomic studies are, therefore, likely to provide new insights into cellular mechanisms involved in cardiac dysfunction and may also provide new diagnostic and therapeutic markers. *Methods*: A literature search from 1991 to 2005 was carried out using the ScienceDirect, and Medline electronic databases. Twenty research studies were identified. These were further examined and synthesized. *Results*: The current battery of proteomic technologies make it possible to characterize global alterations in protein expression associated with processes of human disease. The combined data from microarray and proteomic screens is beginning to establish trends in gene expression, protein modification, and protein movement associated with DCM, hypertrophic cardiomyopathy (HCM), IHD, heart failure and ischemic preconditioning. One such trend, supported by genomics and proteomics, involves the down-regulation of genes and proteins associated with mitochondria and fatty acid metabolism. Proteomic evidence of a strong induction of the enzyme UCH is a provocative finding, that may account for some of the observed increase in proteolytic activity in DCM. Other trends, that are perhaps predictable from previous work, include changes in the expression of myofilament proteins and stress pathways, the preferential increase of protein synthesis genes in HCM over DCM and the contributions of extracellular matrix, cytokines, and inflammation mediators in IHD, infarction and remodelling. Where the data is available, genomic and proteomic screens have also confirmed contributions of apoptosis in hypertrophy, ischemia, DCM and failure. Heat shock proteins are implicated, but may increase or decrease. *Conclusion*: Although the application of proteomics to human heart disease is in its infancy, it is already clear from studies of dilated cardiomyopathy, both in human patients and appropriate models, that it complements traditional candidate gene/protein studies and global genomic approaches and promises to provide new insights into the cellular mechanisms involved in cardiac dysfunction. An additional benefit of these proteomic studies should be the discovery of new diagnostic and/or prognostic biomarkers and the identification of potential drug targets for the development of new therapeutic approaches for combating heart disease.

**44**
**AUXILIARY ALGORITHMS FOR DRUG-TARGET SELECTION**

Stavros J. Hamodrakas

Department of Cell Biology and Biophysics, University of Athens, Athens 157 01, Greece

Our laboratory develops genome-wide application algorithms for prediction of the structure and function of membrane proteins from sequence. The principles, methods and success rates of two predictors of structure and topology of beta-barrel outer-membrane proteins, PRED-TMBB and ConBBPRED, as well as of a predictor of GPCRs, PRED-GPCR, recently developed, are presented. A database of G-proteins, gpDB, is also presented, together with the algorithm PRED-COUPLE that predicts the coupling specificity of an orphan GPCR to G-proteins. An algorithm that depicts, in a versatile way, the structure of membrane proteins, TMRPres2D, is additionally shown. Instructions are given regarding how these tools may help in automatic drug-target selection in genome projects.

**45**
**QUANTITATIVE PROTEOMICS OF PROCHLOROCOCCUS MARINUS**

Kyriacos C. Leptos[1], Jacob D. Jaffe[1,3], Erik Zinser[2], Debbie Lindell[2], Sallie W. Chisholm[2] and George M. Church[1]

[1]Department of Genetics, Harvard Medical School, Boston, MA;
[2]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, U.S.A.;
[3]Present address: The Broad Institute, Cambridge, MA, U.S.A.

Performing whole-cell protein quantitation analysis has proven to be essential in the field of systems biology. For this purpose, we developed MapQuant, a platform-independent open-source software (http://arep.med.harvard.edu/mapquant.html) which, given large amounts of mass-spectrometry data, outputs quantitation for any organic species in the sample. We have previously applied MapQuant in the study of standardization samples at different concentrations on both LCQ and LTQ-FT spectrometers, where we showed linearity of signal with respect to the quantity of protein introduced and acceptable variance among multiple injections. Additionally, ionization suppression issues using protein mixtures of medium complexity at different concentrations were addressed.

The *Prochlorococcus* species is an abundant marine cyanobacterium that contributes significantly to the primary production of the ocean and whose life cycle is synchronized to the solar day (the "diel cycle"). MapQuant was used to quantify the proteins in a time-series data set that includes 25 time-points distributed along a 48-hour period (two diel cycles) of the strain MED4 of *Prochlorococcus marinus*. The data acquisition of the trypsinized proteome took place on a Finnigan LTQ-FT mass spectrometer using short gradients. MS/MS spectra were interpreted using the program SEQUEST. MapQuant algorithms include morphological operations, noise filtering, watershed segmentation, peak finding and fitting, peak clustering and de-isotoping.

MapQuant outputs a list of potential organic species, by reporting four physical attributes for each isotopic cluster that it deconvolves (MapQuant Isotopic Clusters or MQICs). These amount to an average of 17,000 organic species per run. However, the number of peptides identified through MS/MS spectra is limited (only 3%). For this reason SEQUEST IDs from the same organism were utilized, which had been identified in past large-scale LC/LC/MS/MS experiments, to develop an algorithm that will assign their identities to the remaining unassigned number of MQICs (97%).

Our end goal is to be able to perform quantitation for most peptides found in the 25 time-points of the two diel cycles and to understand how carbon fixation, light-response and cell division are coordinated throughout the daily cycle.

**46**
**MACRO-EFFECT OF MICRO-RNA ON HUMAN HEALTH**

Bernard Michot, David Vilanova, , Olivier Delfour and Vera Atzorn

CNRS, France

Non-coding RNA constitute a highly integrated functional network playing key roles in gene expression. Variation in this regulatory architecture is at least equally, if not more, important than variation in the protein components. They may determine the differences between individuals and species, including susceptibility to diseases. This is underlined by recent experimental data demonstrating the direct impact of microRNA function on human health. The functional assessment of every gene is an ongoing challenge with particular importance for the comprehension of human disease and the discovery of novel, valid diagnosis, prognosis and drug targets. Importantly, genetic regulatory network data derived from analysis of the transcriptome activity is substantially incomplete if microRNA sequences are not included in the interrogating set. Unfortunately, the vast majority of microRNA is tissue-specific and temporally expressed at low levels, thus making it difficult to detect with conventional experimental methods. What we know about non-coding RNA today is only the tip of an iceberg.

Therefore, new methods for high-throughput characterization of gene function are needed. This challenge was addressed and a novel, powerful discovery platform developed, that allows identification and validation of the to date hidden regulatory networks implying non-coding RNA. Our technology and expertise offer solutions for the use of this emerging class of molecules of therapeutic interest.

**47**

**SYSTEMS BIOLOGY OF *SPIROPLASMA MELLIFERUM***

Achilleas Frangakis

EMBL, Germany

Evidence has recently accumulated that not only eukaryotes, but also bacteria, can have a cytoskeleton. We used cryo-electron tomography to study the three-dimensional (3-D) structure of *Spiroplasma melliferum* cells in a close-to-native state at ~4 nm resolution. These cells were shown to possess two types of filaments arranged in three parallel ribbons underneath the cell membrane. These two filamentous structures are built of the fibril protein and the actin-like protein MreB. Based on our structural data, we could model the motility modes of these cells and explain how helical Mollicutes can propel themselves by means of coordinated length changes of their cytoskeletal ribbons.

**48**

**GENOME-BASED TRANSCRIPTIONAL PROFILING AND BIOINFORMATIC APPROACHES TO UNDERSTAND ECTODERMAL ORGAN FORMATION**

Lakshmi Raj, Erin Howe, Jinyun Chen and Marianna Bei

Cutaneous Biology Research Center, Massachusetts General Hospital and Harvard Medical School, Boston, U.S.A.

Our studies in the past have shown that one class of genes, the *Msx genes*, controls the formation of several organs, including that of epithelial appendages during mouse development. In particular, epistasis and functional analysis, using the genetically engineered *Msx1* and/or *Msx2* mouse mutants, revealed an *Msx*-controlled genetic hierarchy, where several families of growth and transcription factors are involved (FGFs, BMPs, LEF-1, DLX1, DLX2, RUNX2). We have also studied an MSX nonsense human mutation in a family affected by an Ectodermal Dysplasia Disorder (Ectodermal Dysplasia Disorders are groups of heterogeneous disorders characterized by defects in at least two ectodermally-derived organs such as teeth, nails, hair and sweat glands) known as Witkop Syndrome. An important finding to emerge from these studies is that early formation of many mammalian organs appears to be controlled by conserved regulatory gene-cascades. The extent, however, to which pathways are common between organs is still preliminary. Moreover, the vast majority of genes involved and the molecular mechanism by which these molecules interact to organize a pattern formation are still unknown.

In that context, using recent advances in genomics and high-throughput technologies, we are currently involved in the identification of novel genes and pathways whose function is important to ectoderm-derived organs. We identified: (i) known genes with established functions, (ii) previously characterized genes with so far unknown function in the developing ectodermal organs and (iii) genes of unknown function.

**49**

**MICROARRAY PHENOTYPE CLASSIFICATION USING ENSEMBLE MACHINE LEARNING**

Andrei Dragomir, Giannis Maraziotis and Anastasios Bezerianos

Department of Medical Physics, Medical School, University of Patras, Rio 26500, Greece;
e-mail: {adragomir, imarazi}@heart.med.upatras.gr, bezer@patreas.upatras.gr

Molecular characterization of different tumor types using gene expression profiling is expected to uncover fundamental aspects related to cancer diagnosis and drug discovery. There is, therefore, a need for reliable, accurate sample classification tools, capable at the same time of providing meaningful gene information. Generally, a class prediction task consists of employing a subset of the data with known classification (training set) to derive a decision function able to perform class discrimination. The decision function is subsequently used to make class assignments to new patterns. However, most of the classification methods face difficulties, either due to the classical "curse of dimensionality" (which plagues most classifiers that attempt to determine the decision boundaries in a high-dimensional feature space) or the related overfitting problem. Specific problems related to microarray data, such as noise corrupted datasets and large variability, pose supplementary challenges during the analysis. Consequently, these issues are addressed by means of an ensemble approach, based on Support Vector Machines (SVM) and Boosting.

Our approach of constructing an ensemble is that of an additive model, which predicts the class label of a new data point by performing a weighted sum of a set of component classifiers in such a way that the weighted sum fits the data well. Boosting is an efficient and flexible method based on the this principle. Specifically, it incrementally adds one new

classifier at a time to an ensemble. Each new classifier is constructed by a learning algorithm that tries to minimize the classification error on a weighted training dataset. At each iteration step, the current classifier-weighted error is applied to update the weights of the training examples. The desired effect is to place more weight on the training examples that were misclassified and less weight on examples that were correctly classified. Therefore, in subsequent iteration, the boosting framework constructs progressively more difficult learning problems. Intuitively, we can imagine that subsequent classifiers concentrate on class boundary regions of the data space, where classification decisions are difficult to make. An extension for the multi-class classification is also suggested.

**50**
**FULLY AUTOMATED WORKFLOW FOR THE ANALYSIS OF PROTEIN EXPRESSION LIBRARIES BY MALDI-MS**

S. Bailey[1], A. Wattenberg[1], S. Felske-Müller[1], A. Lücking[2], C. Gutjahr[2], K. Schulte[2], V. Gruß[1], C. Scheer[1], S. Müllner[1] and M. Blüggel[1]

[1]Protagen AG, Emil-Figge-Str. 76A, D-44227 Dortmund; [2]Ruhr-University Bochum, Universitätsstr. 150, D-44801 Bochum, Germany

For protein biochip production, we use our large-scale expression library resulting in over 10,000 recombinant unique human proteins (Uniclone® library). To monitor the production of the protein content prior to protein spotting, a highly automated workflow was used to identify and verify the expressed proteins by peptide mass fingerprint (PMF) and peptide fragmentation fingerprint (PFF). The workflow includes robots for spot picking, protein digestion and target preparation, the Ultraflex TOF-TOF mass spectrometer and ProteinScape, a database for proteome data management (Bruker Daltonik GmbH). These instruments allow the high throughput from hundreds to thousands of samples with minimal assignment of personnel. *Methods*: A first step in quality control is one-dimensional electrophoretic separation, to verify the molecular weight and the homogeneity of the recombinant protein. From these gels, pieces are cut using a spot picking robot (PROTEINEER spII). The in-gel digest and the target preparation is done by the digestion robot (PROTEINEER dp). The MALDI-MS and –MS/MS spectra are acquired in a fully automated way by the Ultraflex TOF-TOF. The ensuing database searches are started automatically at the import of the processed spectra into ProteinScape. The data management from instrument to instrument (spot picker – digester – MALDI-MS – ProteinScape) is provided by transponder technology. *Results*: We analyzed 672 samples within 5 days using the described fully automated workflow. Over 400 protein samples could be identified directly by either peptide mass fingerprint (MS) or by peptide fragmentation fingerprint (MS/MS), or both. One major cause for unsuccessful automated protein identification was due to the low molecular weight of some of the recombinant proteins (< 10 kDa). In these cases, data acquisition and automated data interpretation had to be optimized for PFF since PMF is not applicable. *Conclusion*: A fully automated workflow for quality control of large-scale protein expression libraries was employed successfully. The next step is to analyze thousands of samples derived from our Uniclone® library.

**51**
**INTEGRATED LC-ESI-MALDI-MS/MS ON A DISPOSABLE MALDI TARGET PLATFORM**

Stephanie Hahner[1], Markus Lubeck[1], Martin Schuerenberg[1], Christian Ziegmann[2], Wolfgang Jabs[1] and Detlev Suckau[1]

[1]Bruker Daltonik, Bremen; [2]Eppendorf AG, Hamburg, Germany

The sensitive detection of PTMs from protein digests requires achieving high protein sequence coverage, a high sensitivity of the detection system and for MALDI-TOF a low probability of cross-contaminations between LC-MALDI runs for the detection of low abundant peptide species. In an integrated approach, a capLC system (Agilent 1100) was coupled using a 1:1 post column split to a fast scanning ESI-ion trap mass spectrometer and to a robot for fraction spotting onto a MALDI target. After real-time LC-MS/MS analysis on the trap and database searching, off-line LC-MALDI-TOF/TOF analyses were performed in a data-dependent fashion. MALDI was used: i) to provide additional peptide identifications that were not observed in previous ESI analyses; ii) to validate unsafe ESI matches, under conditions of full automation.

The goal of these experiments was to approach 100% sequence coverage by MS and MS/MS from isolated proteins, such as the ones from 2-D gels. Such a level of detail is required for unambiguous identification of splice variants, modifications (PTMs) or polymorphisms or other sequence variations in proteins. It is, therefore, essential for detailed protein structure elucidation. As several PTMs such as sulfation or phosphorylation can be present at sub-stoichiometric levels, the detection of low abundant peaks for the downstream acquisition of MS/MS spectra must safely exclude the possibility of MALDI target contaminations from previous runs. Disposable AnchorChip MALDI targets (Bruker) were used here. In addition, prespotted α-cyano-4-hydroxycinnamic acid matrix anchors,

(700 μm) allowed the direct application of the LC-eluate onto the target. One hundred amol detection sensitivity was achieved on these matrix anchors, as they concentrate the analyte from the larger droplets that were deposited.

Data-dependent MALDI-MS/MS analysis subsequent to online-ESI-MS/MS provided a significant increase in sequence coverage. As an LC-run is immobilized on a MALDI target, the deposited fractions were available for re-analysis even after several days, if careful data analysis suggested some to be particularly important to the final result. The disposable targets are particularly useful for extensive sample archival.

**52**

## MASS SPECTROMETRIC SEQUENCING OF PROTEIN-DERIVED CARBOHYDRATES USING MALDI-TOF/TOF MASS SPECTROMETRY

Urs Lewandrowski[1], Anja Resemann[2], Detlev Suckau[2], Lars Vorwerg[2], Herbert Thiele[2] and Albert Sickmann[1]

[1]Rudolf-Virchow Center for Experimental Biomedicine, Wuerzburg; [2]Bruker Daltonics, Bremen, Germany

Among the most abundant post-translational modifications, glycosylations comprise a highly complex group of carbohydrate modifications to the polypeptide backbone of proteins. The two most common forms, N- and O-glycosylation, are both characterized by branched three-dimensional structures which are highly diverse in form and size. Upon appearance of various mass spectrometric techniques from the proteomics sector, glycan analysis also profits from the recent advances in instrumental design. MALDI-TOF/TOF sequencing of protein-derived carbohydrates is described in respect to various matrices and system parameters.

Oligosaccharides were obtained from various sources; some model compounds (Sigma) and some obtained by glycosidase treatment from glycoproteins. Standard dried droplet preparation on stainless steel MALDI targets was employed; DHB was the typical matrix but other matrices, such as sDHB and THAP, were also evaluated. MS and LIFT-TOF/TOF MS/MS spectra were obtained on an Ultraflex TOF/TOF (Bruker Daltonics) either under conditions of unimolecular decomposition (LID) or high energy CID (heCID), primarily using Ar as the collision gas. MS/MS spectra analysis was supported by new software tools.

HeCID resulted in an enhanced occurrence of analytically valuable cross-ring fragmentations (A- and X- ion series) in addition to the commonly observed interglycosidic cleavages (B- and Y-ions). This offers the opportunity of fast initial sequencing of carbohydrate structures with the possibility of differentiating between structural linkage isomers. During the experimental approach, the influence of various system parameters was examined, including different matrices, LID *vs*. heCID and sample purity. It turned out that complex carbohydrate samples resulted in significantly different MS/MS spectra as compared to highly purified oligosaccharides, possibly due to significantly increased laser fluence settings. Another important aspect in this work was the development of an analytic strategy that included sample preparation, the fragmentation regime and spectra interpretation, as well as structure visualization.

**53**

## SCREENING FOR SEQUENCE ABERRATIONS AND MODIFICATIONS ON THE PROTEOME LEVEL

Peter Hufnagel, Ulrike Schweiger-Hufnagel, Sven Brand, Marina Behrens, Carsten Baessmann, Detlev Suckau and Arnd Ingendoh

Bruker Daltonik, Bremen, Germany

2-D PAGE followed by mass spectrometry is an outstandingly useful technique, as the detailed protein structure can be elucidated on the level of each individual spot. Even more information can be obtained from the combination of complementary MS techniques. Then, various features can be illuminated, as shown in this approach. This is facilitated by an integrated software environment that comprehensively organises information on all analysis levels.

Total protein was separated on a 2-D gel. After colloidal coomassie staining, the gel was mechanically fixed using a frame equipped with a transponder. The protein spots were excised and transferred to microtitre plates, which were protected against dust and evaporation in dedicated, transponder-equipped plate holders. There, the proteins were digested by porcine trypsin. For high-throughput analyses, a MALDI -TOF/TOF instrument was used. ESI Ion trap analysis was performed subsequent to nano LC separation (75 μm column). Electronic transponders allowed tracking of the complete sample information and status throughout the whole process.

A three-step strategy was employed in a fully automatic workflow ("WARP"): (i) Peptide Mass Fingerprints (PMFs) were acquired first and based on the results. (ii) MALDI-TOF/TOF was used to verify identifications and increase the identification yields using MS/MS analyses. This strategy was also extended to the targeted analysis of peptides that remained unrelated to identified proteins. That approach allowed targeting of unexpected structural features such as modification or mutations. (iii) For all unidentified or uncertain spots, a more powerful yet more time-consuming technique was used: nano LC-MS/MS.

Resulting in high sequence coverage, nano LC-MS/MS was the method of choice for most detailed structural

characterisation. Unexplained LC-MS/MS compounds were submitted to automatic *de novo* sequencing and internal homology searches for the detection of sequence aberrations and posttranslational modifications.

**54**
**ISD-MALDI AND T3-SEQUENCING FOR TOP-DOWN SEQUENCE ANALYSIS OF RECOMBINANT PROTEINS: APPLICATIONS AND TECHNICAL ASPECTS**

Anja Resemann, Detlev Suckau and Arnd Ingendoh

Bruker Daltonik, Bremen, Germany

Quality control of recombinant proteins is important for a growing number of applications, including the production of protein arrays. Simple, cost-efficient and fast methods are required to screen such recombinant proteins for the proper primary structure, terminal processing and any type of modifications. Of the classical methods, N-terminal sequencing is sensitive but slow and unsuitable for proteins with N-terminal modifications; chemical approaches for C-terminal sequencing are impractical for routine applications; protein mass fingerprints frequently do not provide information about the termini.

Top-down protein analysis, *i.e.* direct MS/MS analysis of intact proteins without proteolytic digestion, can overcome these obstacles. A series of N- and C-terminal "peptides" are generated by in-source-decay (ISD) in the MALDI-TOF spectrometer from proteins as large as 100kD. ISD sequence tags are typically 20-150 residues long, commencing at *ca.* 7-10 residues from the termini of the protein. In most cases, this information is sufficient to characterize the termini of the protein and indicate the presence of N-terminal modifications. More detailed information on the N- and C-termini can be obtained in a second step, by MS/MS analysis of the terminal peptides in the TOF/TOF part of the instrument.

(i) We evaluated ISD-MALDI-based N- and C-terminal characterization for a variety of "real-life" proteins, which are often heterogeneous and are delivered at low concentrations or in MS incompatible buffers. Examples include N-terminally blocked proteins and proteins that underwent partial proteolysis. (ii) The use of hydrophobic magnetic nano-particles and prestructured MALDI targets for sample preparation is essential for optimal ISD performance and sequence coverage. (iii) Examples for MS/MS analysis of terminal fragment ions are presented, that provide an exhaustive characterization of the terminal residues not covered by the direct ISD experiment. (iv) The potential of linear-mode ISD for complete sequencing of medium-sized proteins up to 20kD is discussed.

A sensitivity level of 500 fmol was achieved, which is expected to open up the application scope of top-down analysis to the field of proteins purified from natural sources and even top-down proteomics.

**55**
**ENRICHMENT OF GLYCOPROTEINS AND GLYCOPEPTIDES SUPPORTED BY MAGNETIC PARTICLES AND DETECTED BY MALDI-TOF-MS**

Katrin Sparbier, Sonja Koch, Irina Keßler, Thomas Wenzel, Markus Kostrzewa and Arnd Ingendoh

Bruker Daltonik, Bremen, Germany

Glycosylation is, by at least 50 to 80 percent, the most common form of post-translational modification of proteins. Since the degree and the type of glycosylation depends on the *status quo* of cells and is mostly linked to certain diseases, the discovery and identification of those modified peptides and proteins gain increasing importance in diagnosis. In particular, aberrant or missing glycosylation can be utilized as a profiling parameter. MALDI-TOF mass spectrometry is an ideal technique for identifying a large number of peptides and proteins and their corresponding modifications. Enrichment, isolation and purification of glycosylated peptides and proteins from different sources – a prerequisite for MALDI-TOF mass spectrometry – can be attained by affinity chromatography supported by magnetic particles.

Concanavalin A (ConA)- and di-boronic acid-functionalized magnetic particles were developed to enrich glycosylated peptides and proteins. ConA specifically binds the mannosyl and glucosyl residues of polysaccharides and glycoproteins containing unmodified hydroxyl residues at positions C3, C4 and C6 and can be applied as a general tool for capturing N-glycosylated peptides and proteins. In contrast, boronic acid forms a covalent bond with the 1,2-cis-diol group containing molecules like mannose, galactose or glucose. Thereby, it additionally facilitates the enrichment of the more heterogeneous O-linked oligosaccharides, but not the capture of N-linked oligosaccharides of the complex type.

The functionalized beads were employed to establish and optimize protocols for the binding and detection of glycosylated peptides and proteins with respect to an automated workflow and the subsequent detection and identification by MALDI-TOF mass spectrometry. For several model proteins, the capture could be demonstrated by SDS-PAGE and MALDI-TOF mass spectrometry. According to the type of glycosylation, high mannose, hybrid or complex type, the different proteins were enriched by ConA- or boronic acid-functionalized beads,

respectively. RNase B could be isolated from spiked human serum samples by ConA beads according to the well-known N-linked high mannose oligosaccharide structure.

Biochemistry *13(15)*: 3124-3130, 1974.

**56**

## IDENTIFICATION OF SERUM PEPTIDES BY MAGNETIC BEAD-BASED FRACTIONATION COMBINED WITH LC-MALDI-TOF/TOF

Thomas Elssner[1], Dagmar Niemeyer[1], Ralf Ketterlinus[2], Detlev Suckau[2], Stephanie Hahner[2], Sören-Oliver Deininger[1], Markus Kostrzewa[1] and Arnd Ingendoh[2]

[1]Bruker Daltonik, Leipzig; [2]Bruker Daltonik, Bremen, Germany

The discovery and identification of naturally occurring peptides and proteins reflecting the *status quo* in biochemical mechanisms and pathways are gaining increasing importance in the context of diagnosis of diseases (*e.g.* cancer). MALDI-TOF mass spectrometry is a unique tool for the detection of such peptides within complex profile spectra and their subsequent identification. Fractionation, enrichment and purification of peptides and proteins from complex biological materials necessary for MALDI-TOF analysis is mandatory and can be achieved by magnetic bead-based chromatography. Identification of the discovered peptides can be directly accomplished by MALDI-TOF/TOF analysis, whereas larger proteins may not be identified using this approach, since protein digestion is required to identify them.

A new strategy combining magnetic bead-based chromatography and LC-MALDI-TOF/TOF analysis allows the further specific enrichment of peptides and proteins and the identification of a larger number of molecules within the low molecular mass range by MALDI-TOF/TOF analysis. Various magnetic beads with different functionalities (*e.g.* hydrophobic, cation exchange) were used for peptide and protein capturing from human serum samples. One part of the eluate recovered from the magnetic beads was used for direct acquisition of profile spectra and the other part was analyzed on a capLC system. The eluate was spotted into discrete spots on a MALDI target. All sample fractions were analyzed automatically by MALDI-TOF/TOF mass spectrometry. Database searching allowed the identification of all peptides with molecular weights of less than 3000 Da.

In the presented approach, peptide markers were discovered in MALDI MS profiles obtained from magnetic beads at high sample throughput, while their identification was achieved from a low number of LC-runs using TOF/TOF measurements.

**57**

## SAMPLE SONICATION AFTER TRICHLOROACETIC ACID (TCA) PRECIPITATION INCREASES PROTEIN RECOVERY AND RESOLUTION IN TWO-DIMENSIONAL GEL ELECTROPHORESIS

Bruno Manadas and Carlos B. Duarte

Center for Neuroscience and Cell Biology and Department of Zoology, University of Coimbra, Portugal

The resolution of thousands of proteins in 2D-SDS-PAGE is a great task requiring fractionation and building of cybergels from zoom IEF gels. Although improved resolution is achieved with these gels, a higher degree of reproducibility is also required in order to allow quick matching of the gels by the algorithms available and to speed up the analysis.

Protein samples to be analyzed by two-dimensional gel electrophoresis should be free from salts and other compounds that interfere mainly with IEF, such as nucleic acids and lipids. This may be achieved by various means, including TCA precipitation followed by acetone washing. The main problem associated with TCA precipitation concerns solubilization of the pellet. In order to improve protein recovery and to decrease the amount of proteins trapped in the insoluble fraction after TCA precipitation, we have tested the effect of sonication after elution of proteins in a sample buffer containing urea, thiourea, CHAPS and DTT. The effect of sonication was tested on TCA-precipitated protein pellets from hippocampal soluble and membrane/membrane-associated proteins. Sonication decreased the size of the pellets comprising insoluble material, particularly in the latter fraction. Quantification of the total protein content in sonicated and non-sonicated membrane fractions showed that this physical treatment increases total protein recovery by 140%, starting with the same amount of sample. In order to determine the effect of sonication on 2D-PAGE, sonicated and non-sonicated samples were subjected to IEF and SDS-PAGE. Although the pattern of both gels was similar, when silver-stained, gels from sonicated samples showed several new spots and a different focusing pattern of some spots. The effect on sonication on the reproducibility between gels was tested in three replicate gels from non-sonicated and sonicated samples (membrane fractions), radiolabelled with $^{35}$S-amino acids, analyzed with PDQuest™. It was found that higher matching ratios were obtained when gels from sonicated samples were compared with each other than when non-sonicated gels were compared with each other. Taken together, the results show that sonication increases protein recovery and reproducibility between gels, thereby decreasing the time spent in manual editing.

**58**

## TARGETED NANO-LC/MS/MS USING MULTIPLE REACTION MONITORING DETECTION IMPROVES IDENTIFICATION OF LOW ABUNDANCE PROTEINS IN COMPLEX MIXTURES

Christof Lenz[1], Volker Kruft[1] and Henning Urlaub[2]

[1]Applied Biosystems, Frankfurter Strasse 129 B, 64293 Darmstadt;
[2]Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Goettingen, Germany

The reliability of LC/MS/MS protein identification after protease digestion is limited by the number of digest peptides that are detected and selected for MS/MS analysis. One of the major challenges here is the detection of low abundance peptide precursors against the chemical background in MS. Several approaches to circumvent this problem have been proposed, including preferential detection of multiply-charged precursors and the use of precursor ion scans for immonium ions typical of certain amino acids. Using a hybrid triple quadrupole/linear ion trap mass spectrometer, we demonstrated that the specific detection of digest peptides derived from *in silico* digestion of suspected proteins by Multiple Reaction Monitoring (MRM) significantly improves the sensitivity and sequence coverage for the detection of these proteins. *Methods*: Complex mixtures of human spliceosomal proteins were trypsinized. Digest peptides were separated and mass analyzed on a nanoLC system coupled to a hybrid triple-quadrupole/linear ion trap mass spectrometer (4000 Q TRAP™). Tentatively identified or predicted low-abundance proteins were digested *in silico*. MS/MS fragmentation reactions of the resulting peptides expected to produce abundant and specific product ions were translated into MRM transitions monitored throughout the LC/MS/MS analysis. The signal intensity in MRM was used to trigger full product ion spectra using linear ion trap detection. The resulting MS/MS spectra were matched against an NCBInr database using MASCOT® software. *Results*: Seven proteins, that were weakly identified or only suspected in human spliceosomal complexes A and B, were trypsinized *in silico*. The hypothetical digest peptides were subjected to MRM-triggered nanoLC/MS/MS on a hybrid triple quadrupole/linear ion trap mass spectrometer. Comparison to regular MS-triggered nanoLC/MS/MS analysis, both on this system and on a state-of-the-art QqToF instrument, showed significantly improved detection of digest peptides both by MRM and the triggered linear ion trap full product ion spectra, with an average increase of sequence coverage by a factor of 3.4.

**59**

## IMAGE ANALYSIS OF TWO-DIMENSIONAL POLYACRYLAMIDE GEL ELECTROPHORESIS. AN APPROACH TO INCREASE THROUGHPUT

N. Tsokanas, P.Cutler and N.A Jones

Disease and Biomarker Proteomics, GlaxoSmithKline, Stevenage, U.K.

Over recent years, proteomics studies have gained increasing attention from major pharmaceutical companies. The 2-D PAGE approach remains the main technology for study of the proteome. The 2-D gels are ordinarily scanned at a resolution of 100 μm. Sophisticated software(s) have been developed for the analysis of 2-D gels and the quantification of changes in protein expression. However, because of the complexity of the data structure, accurate spot detection and matching of all spots across the entire dataset by the software within desired time-limits is a challenge. Image and data analysis remain a "bottleneck" in the proteomics workflow. In this study, a variety of methods to analyse digitised 2-D gel images are described, which reduce the time required to complete the image and data analysis process to produce accurate quantitative data. Using a combination of statistical methods, data can be analysed in a more efficient manner. The results show that the use of digitised gels scanned at 200 μm can reduce the processing time of IA by a factor of **four** without any loss of qualitative or quantitative data. The highest number of protein changes was identified by the use of the Student's *t*-test; however, application of multivariate statistical methods allows further protein expression changes to be identified

**60**

## CHARACTERISATION OF COMPLEX PROTEIN GEL BANDS BY OFFLINE LC-MALDI QIT TOF MS

Emmanuel Raptakis[1], Rachel L Martin[1] and Willy Bienvenue[2]

[1]Shimadzu Biotech, Manchester, U.K.;
[2]IDE-ALP Pharma, Lyon, France

Traditionally, the identification of complex protein mixtures has required a series of time-consuming protocols, including gel electrophoresis and mass spectrometry-based peptide mass fingerprinting (PMF). Limitations can arise, including co-migration of more than one protein, suppression effects during ionization, or the PMF not correlating with any protein in the databases. Additionally, MS/MS can provide sufficient information to generate useful sequence data. However, when peptides have undergone modifications, $MS^n$ techniques may be required.

Here, in an attempt to circumnavigate many of these issues, digested gel samples were analyzed using LC MALDI-QIT-TOF-MS. The proteins were separated using 1-D gel electrophoresis; bands or spots were excised and digested with trypsin. The digests were separated by micro-LC using a C18 column (10 cm x 0.3 mm ID) and the eluent spotted onto a MALDI target. The flow rates used were 5-50 μl/min. The spotting process was performed by an automated system (AccuSpot, Shimadzu Biotech) which can accommodate up to nine 384-well stainless steel MALDI targets. The matrix (dihydroxybenzioc acid (DHB)) is coaxially introduced to the eluent allowing thorough mixing before deposition on the target. An Axima-QIT (Shimadzu Biotech) was used to analyze the separated tryptic peptides in positive ion mode. Helium was used as the ion trap buffer gas and argon as the collision gas.

Human proteins from BJAB cell full lysate separated on 1-D electrophoresis were analyzed. A band of interest was situated at 33-42kDa. After LC separation and spotting directly onto a MALDI target, the sample was analyzed by the MALDI-QIT-TOF-MS under automation, initially in positive ion MS mode to provide confirmation of the position of peptides on the target and their masses. Next, MS/MS was performed on the peptides located. The resultant fragmentation spectra were submitted to the Mascot search engine (Matrix Science) in order to identify the proteins from which the peptides originated. If the search was unsuccessful, *de novo* sequence information obtained from the $MS^n$ data was then generated. The high mass accuracy and high resolution in MS/MS mode of more than 8000 enabled the confident identification of proteins by both MS/MS database searching and interrogation of sequence tags.

**61**

## PROTEOME PROFILING OF THE MOUSE BRAIN USING MULTIDIMENSIONAL CHROMATOGRAPHY COMBINED WITH OFFLINE LC-MALDI AND ONLINE LC-ESI TANDEM MASS SPECTROMETRY

Spiros D. Garbis, Elena P. Xixi, Konstantinos N. Vougas, Ploumisti Dimitraki, Panagiotis G. Zerefos and Michael Fountoulakis

Biomedical Foundation of the Academy of Athens, Athens, Greece

Global protein analysis in neuroscience, despite the latest advancements in mass spectrometry-based proteomics, is still in its infancy. The key reasons for this insufficiency in identified proteins are the lack of selectivity and sensitivity of the 2-D electrophoresis-based MS methods. In order to expand the application base and sensitivity of the mass spectrometric analysis of the mouse brain proteome, unique protein extraction and enrichment protocols combined with multi-dimensional chromatography (LCn)-tandem mass spectrometry (MS-MS) techniques were used. *Methods*: Whole brain tissue samples were liquid phase extracted for hydrophilic proteins and hydrophobic proteins with detergent and chaotrope-free sodium carbonate and ammonium bicarbonate solutions. The hydrophilic portions were fractionated with a polySULFOETHYL Aspatamide, 4.6 mm ID x 150 mm L x 1000 Angstrom pore strong cation exchange-hydrophilic interaction (SCX-HILIC) HPLC column and further fractionated with a C4, 12 mm ID x 150 mm L x 300 Angstrom pore size, reverse phase column The hydrophobic portion was fractionated with a polyHYDROXYETHYL Aspartamide 4.6 mm ID x 150 mm L x 300 Angstrom pore HILIC HPLC column. All fractionations were performed with the P680 HPLC system(Dionex Corp.). One set of the resulting samples were applied to a 384-spot ground steel target and subjected to matrix-assisted laser desorption time-of-flight MS-MS (MALDI-TOF-TOF-MS-MS)analysis (Ultraflex model, Bruker-Daltonics Corp.). The duplicate set of HPLC fractions were then analyzed with LC-MS-MS. Tandem mass spectrometric analysis of the nano-LC eluent was performed with a nano-electrospray ionization-quadrupule ion trap system (Model SL MSD, Agilent Corp.) and a nano-electrospray hybrid quadrupule time-of-flight (QqTOF) system (QSTAR XL, Applied Biosystems). The study involved comparative investigation by identifying proteins with LC-MALDI-MS and LC-ESI-MS approaches.

**62**

## PROTEOMIC ANALYSIS OF MOUSE BRAIN FOLLOWING ENRICHMENT BY PREPARATIVE ELECTROPHORESIS

Elena Xixi, Ploumisti Dimitraki, Kostantinos Vougas, Sofia Kossida and Michael Fountoulakis

Division of Biotechnology, Foundation for Biomedical Research of the Academy of Athens, Greece

Proteomics is a powerful technology for studying the identity and levels of brain proteins. Changes of protein levels, as well as modifications that occur in neurological disorders, may be informative of the pathogenesis of these disorders and could result in the identification of potential drug targets and markers. To increase the capability of characterizing complex protein profiles, protein mixtures should be separated into simpler fractions, thus increasing the likelihood of detecting low-abundance proteins. Considering that low-abundance proteins are thought to be involved in important biological processes in the brain, identification of these low copy number gene products appears to be a scientific challenge. In this work, proteomic analysis of adult mouse brain tissue was performed

following enrichment by preparative cell electrophoresis. This was performed using the PrepCell apparatus in the presence of 0.1% lithium dodecyl sulfate. Samples were electrophoresed through a cylindrical polyacrylamide gel and were fractionated according to their molecular masses. The protein content of the fractions collected was first analyzed by SDS-PAGE and then by two-dimensional gel electrophoresis. Protein identification was performed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Analysis in the MS/MS mode was performed for the low molecular mass proteins not identified during the analysis in the MS mode. The present analysis resulted in the identification of 380 different proteins. Among those were heat shock proteins, enzymes, transport proteins, transcriptional activators, signal transduction molecules as well as proteins with a number of other functions. Preparative electrophoresis may be a useful method in the investigation of the brain.

**63**
**PROTEOMIC ANALYSIS OF AMNIOTIC FLUID SUPERNATANT SAMPLES IN PREGNANCIES WITH DOWN SYNDROME**

G.T. Tsangaris[1], P. M. Karamessinis[1], A. Kolialexi[2], S.D. Garbis[1], A. Antsaklis[3], A. Mavrou[2] and M. Fountoulakis[1]

[1]Division of Biotechnology, Center of Basic Research, Foundation of Biomedical Research of the Academy of Athens, Athens;
[2]Medical Genetics and [3]1st Department of Obstetrics & Gynecology, Athens University School of Medicine, Athens, Greece

Proteomic analysis, which combines two-dimensional (2-D) electrophoresis and mass spectrometry (MS), nowadays has wide application in biological and medical sciences, mainly for protein screening in tissues obtained from healthy and diseased conditions, for the detection of diagnostic markers and drug targets. Five amniotic fluid supernatant (AFS) samples were obtained from women in the second trimester of pregnancy, known to carry fetuses with Down's syndrome and ten from women carrying chromosomally normal fetuses with informed consent. Samples were analysed by 2-D electrophoresis. Gel comparison showed differences in a mean pool of 26 spots per gel. The indicated spots were excised and proteins were further analysed by peptide mass fingerprint (PMF), Post Source Decay (PSD) with MALDI-MS-MS and nanoelectrospray QqToF-MS-MS with CID. Alpha-1-microglobulin (AMBP_HUMAN, P02760), splicing factor arginine/serine-rich 4 (SFRS4_HUMAN, Q08170), collagen alpha 1(I) chain (CO1A1_HUMAN, P02452), collagen alpha 1(III) chain (CO3A1_HUMAN, P02461), collagen alpha 1(V) chain (CO5A1_HUMAN, P20908) and basement membrane-specific heparan sulfate proteoglycan core protein (PGBM_HUMAN, P98160) were increased in the cases with Down's syndrome as compared to the normal ones, while the insulin-like growth factor binding protein 1 (IBP1_HUMAN, P08833) was decreased. SFRS4, CO1A1, CO3A1, CO5A1 and PGBM were present as fragments. The observed changes were reproducible in all pregnancies with Down's syndrome tested, thus these proteins are possible potential markers useful for prenatal diagnosis.

**64**
**THE NORMAL HUMAN AMNIOTIC FLUID SUPERNATANT PROTEOME**

G.T. Tsangaris[1], A. Kolialexi[2], P.M Karamessinis[1], A. Antsaklis[3], M. Fountoulakis[1] and A. Mavrou[2]

[1]Division of Biotechnology, Center of Basic Research, Foundation of Biomedical Research of the Academy of Athens, Athens;
[2]Medical Genetics and [3]1st Department of Obstetrics & Gynecology, Athens University School of Medicine, Athens, Greece

Proteomic analysis, which combines 2-D electrophoresis and MS, nowadays has wide application in biological and medical sciences, mainly for protein screening in tissues obtained from healthy and diseased conditions, for the detection of drug targets and diagnostic markers. Amniotic fluid samples are routinely used for prenatal diagnosis of a wide range of fetal abnormalities. In prenatal diagnosis, proteomics have been applied for the analysis of tissues from abnormal fetuses in order to detect differences in the protein profile as compared to the normal one and to determine possible diagnostic tools. A detailed two-dimensional protein database for the normal human AF cells, including 380 different gene products, has already been reported. In the present study, the two-dimensional protein database of the normal human AF supernatant was constructed. Ten AF supernatant samples from women carrying normal fetuses were analyzed by 2-D gel electrophoresis. A mean of 412 proteins per gel were analyzed by MALDI-MS and MALDI-MS-MS and the protein identification was carried out by PMF and PSD, respectively. We constructed a 2-D protein map comprising 136 different gene products. The majority of the identified proteins were secreted proteins, carriers and immunoglobulins. Thirteen hypothetical proteins were also included. The normal AFS proteome map is a valuable tool to study aberrant protein expression and in the search for new proteins as possible markers for the prediction of abnormal fetuses.

**65**
**THE HUMAN URINE PROTEOME**

P.G. Zerefos[1], K. Vougas[1], A. Petrolekas[2], K. Stravodimos[2], A. Giannopoulos[2], A. Vlahou[1] and M. Fountoulakis[1]

[1]Division of Biotechnology, Foundation for Biomedical Research of the Academy of Athens; [2]Department of Urology, Laikon Hospital, Medical School of Athens, Greece

Human urine contains metabolites and biomolecules with molecular weights less than 40 kDa, which can easily pass through the kidney filtering system. On the other hand, substances between 40 and 70 kDa are present only in 0.1-1% (w/v), while larger macromolecules are absent under normal conditions. The various components of urinary protein excretion are useful indicators of renal function and of human health in general. Urine samples are easily attainable, thus making them ideal substrates for biomarker research. Consequently urine proteome studies are very important. For this purpose, proteomic analysis of urine samples concentrated by ultra filtration was performed. Samples from healthy individuals were pooled, cleaned by TCA precipitation and analyzed by 2-dimensional electrophoresis (2-DE) on pH 3-10 non-linear IPG strips. Protein spots were excised, trypsinized and analyzed by MALDI-TOF MS. To increase protein detection and identification, we performed preparative electrophoresis using the PrepCell apparatus prior to the 2-DE and MS. Eluates were collected from the PrepCell and analyzed by SDS-PAGE. According to these results, eluates were pooled, desalted by TCA precipitation and analyzed by 2-D electrophoresis. In total, 778 proteins (141 unique) were identified following enrichment by PrepCell, in contrast to 171 (42 unique) in the starting material. Additionally, the median molecular weight of the identified proteins from the preparative electrophoresis was significantly lower in comparison with the proteins of the starting material (Median values: 39886 Da and 71317 Da, respectively), thus enabling the analysis of low molecular weight proteins. In conclusion, the more coherent and detailed study of human urine proteome performed with the proposed methodology may provide the basis for the detection of candidate biomarkers.

**66**
**TOWARDS THE CHARACTERIZATION OF THE ZEBRAFISH PROTEOME**

T. Karinos[1], K. Vougas[2], A. Vlahou[2], M. Fountoulakis[2], G. Koumoundouros[1] and C.N. Flytzanis[1]

[1]Department of Biology, University of Patras; [2]Division of Biotechnology, Foundation for Biomedical Research of the Academy of Athens, Greece

Zebrafish [*Danio rerio*] has been a model organism for some time in the fields of genetics and developmental biology, due to its short generation time, high reproductive potential even in captivity, cost-effective breeding and transparency of developing embryos. In this study, zebrafish were examined by proteomics technologies in an attempt to enrich the online database of known proteins of this species. We analyzed whole protein extracts from muscle and brain tissues by means of 2-D electrophoresis and MALDI-TOF mass spectrometry. The data obtained were blasted using the combined Swiss/TrEMBL and IPI databases. A total of 184 (brain) and 121 (muscle) individual proteins were identified. Of these, 42% and 55%, respectively, were hypothetical or non-characterized proteins. These were then categorized according to their possible function by blasting their amino acid sequence in the above databases. The results showed that most of the hypothetical and non-characterized proteins had a significant homology (70%-90%) with well-characterized proteins from other vertebrate species such as human, mouse, *Xenopus*, medaka fish *etc*. Specifically, most of them were enzymes, or structural molecules, as expected, due to their high abundance in almost every tissue. Enriching the online database of zebrafish will facilitate proteomics studies on this species. Our future plans include the determination of qualitative and quantitative protein differences between males and females, as well as a study of the proteomic profile in developing embryos associated with temperature-sex determination.

**67**
**PROTEOMIC IDENTIFICATION OF DIFFERENTIALLY-EXPRESSED PROTEINS IN RENAL CELL CARCINOMA: EVALUATION OF MEMBRANE PROTEIN EXTRACTION AND ENRICHMENT METHODOLOGIES**

V. Aggelis, R.A. Craven, R. Tonge, P.J. Selby and R.E. Banks

St James University Hospital, School of Medicine , University of Leeds, U.K.

Renal cell carcinoma accounts for 2-3% of diagnosed adult malignancies and is responsible for 95,000 deaths per year worldwide. Furthermore, there have been reports of its increasing incidence and mortality in the last decade in the UK and across the world, especially in females. Currently, the only effective form of treatment is surgery, however 50% of the cases present with locally advanced or metastatic disease that displays great resistance to chemotherapy and radiotherapy. Although alternative biological therapies, such as antibody- and vaccine-based approaches directed at tumour membrane antigens, such as G250 and Her-2/neu, have been tried with promising results, there still remains a

need for the identification of novel disease biomarkers that can be used as prognostic or diagnostic factors, or even constitute potential therapeutic targets. The primary goal of this study was the proteomic identification of protein components of the plasma membrane of renal cancerous cells that could potentially be used as disease biomarkers. Two diverse approaches to extract and enrich membrane proteins, based on cell surface biotinylation and subcellular fractionation, were evaluated by the use of antibodies specific to different subcellular marker proteins. In addition, a comparative study of common commercially available detergents, such as sodium dodecyl sulphate, NP-40, deoxycholate and CHAPS in solubilizing membrane proteins, was also conducted by using 1-D and 2-D PAGE. Through comparative evaluation of these methodologies, we aimed to develop the ultimate robust membrane protein extraction method that will allow the identification of possible membrane renal cancer biomarkers.

**68**
**PROTEIN PROFILES OF HUMAN SALIVA: IDENTIFICATION OF TUMOR-RELATED PROTEINS**

K. Krapfenbauer, D. Turhani, H. Langen, M. Fountoulakis and D. Thurnher

Roche Center for Medical Genomics, Switzerland

Protein analysis of saliva, as a clinical application, offers an attractive, simple and rapid diagnosis tool for the short- and long-term monitoring of pathological disorders and drug therapy. To collect saliva, either in the pure or in the whole form, is a relatively easy and a non-invasive procedure; it is not harmful to the patients and has no complications at all. However, the fluid collection must be clearly defined due to variations in saliva composition and flow-rate. The study discusses the identification of tumor-related proteins in saliva by proteomic analysis, which can be used for the detection and identification of disease markers, specific for oral squamous cell carcinoma (OSCC) in the head and neck. In the frame of our study, matched saliva samples from five control patients and one with suspicion of oral squamous carcinoma were investigated. Approximately 800 spots were identified, corresponding to 117 different gene products. The list of identified proteins included a large number of structural proteins like keratins, keratin subunits, enzymes and enzymes inhibitors, cytokines, immunoglobulins, as well as amylase and other salivary specific glycoproteins. The majority of proteins were localized in the cytosol of oral epithelial cells found as unsolved debris in saliva. Two of them are cell-specific for SCC and may be useful as prognostic "salivary biomarkers", as a means of monitoring general health and in the diagnosis of SCC disease in head and neck.

**69**
**BIOMARKERS OF MYOCARDIAL INJURY**

Michaela Scigelova[1], Tori Richmond[2], Manor Askenazi[2], Jennifer Sutton[2] and Leo Bonilla[2]

Thermo Electron, Hemel Hempstead, U.K.;
BRIMS Center, Thermo Electron, Cambridge, MA, U.S.A.

Myocardial infarction (MI) is a major cause of morbidity and mortality in developed countries worldwide, with approximately 500,000-700,000 deaths caused by ischemic heart disease occurring in the United States alone. The build-up to a massive heart attack is hard to predict, particularly at the early stages when prevention is still possible. When a heart muscle dies, cellular membranes lose integrity and intracellular enzymes and proteins slowly leak into the bloodstream. These enzymes and proteins can be detected by a blood sample analysis. The concentration of enzymes in a blood sample – and more importantly, the changes in concentration found in samples taken over time – correlates with the amount of heart muscle that has died. All existing cardiac markers are detected hours after the onset of myocardial necrosis (irreversible); there is no marker for early ischemia (reversible).

Proteomics – both large-scale protein identification and quantitative differential expression – may be used to identify improved diagnostic markers and pathways triggered during myocardial injury that might, one day, serve as therapeutic targets. The model of alcohol septal ablation, a medical procedure used to reduce a heart hypertrophy after infarction, can be viewed as a "planned" myocardial infarction, and thus serve as a suitable model for the discovery of relevant biomarkers. Preliminary results from this model allowed the quantification and identification of known biomarkers of cardiac disease (*e.g.* C-reactive protein), as well as several biologically relevant low-level soluble factors. Validation of the novel markers in broader patient cohorts using complementary technologies (*e.g.*, SRM scanning using a triple quadrupule mass spectrometer or, when available, immunoassays) is ongoing.

**70**
**PROTEOMIC PROFILING OF CATHEPSIN D DEFICIENT MOUSE MODEL**

M. Lindfors[1], R. Soliymani[2], M. Baumann[2], P. Saftig[3] and J. Tyynelä[1]

[1]Institute of Biomedicine/Biochemistry and Neuroscience Research Program and
[2]Protein Chemistry Unit, Biomedicum Helsinki, P.O. Box 63 00014 University of Helsinki, Finland;

[3]Unit of Molecular Cell Biology and Transgenic Research, University of Kiel, Otto Hahn Platz 9 D24118 Kiel, Germany

The neuronal ceroid-lipofuscinoses (NCLs) are a group of recessively inherited lysosomal storage diseases characterized by progressive neurodegeneration and premature death. At present, defects in nine different proteins are identified as causes of NCLs. There are two naturally occurring and several genetically engineered mouse models of NCL. Of these, cathepsin D (CTSD)-deficient mice have the most severe phenotype. They develop neurological symptoms including tremor and epilepsy, and die at the age of four weeks. Furthermore, they show a number of neuropathological changes and neurodegeneration in the CNS.

The aim of this study was to identify alterations of protein expression in the brains of CTSD knockout mice (CTSD -/-) compared to controls. Information about these alterations might lead to identification of the metabolic pathways underlying the neuronal cell death observed in CTSD -/- mice, and, potentially, in NCL diseases in general.

In the present study, the mouse brain proteome was studied by conventional two-dimensional gel electrophoresis (2-D) combined with MALDI-MS identification of selected proteins. Preliminary results show a number of proteins with at least 2-fold alteration in expression. At present, the identification of selected spots with altered expression is ongoing, and we have identified proteins related to lipid metabolism, signal transduction and synaptic functions. In order to better understand the synaptic functions potentially involved in the pathogenesis, we aim to analyze the protein expression in isolated synaptic vesicles.

**71**

## LEVELS OF MEMBRANE-ASSOCIATED AND mRNA EXPRESSION OF COMPLEMENT REGULATORY PROTEINS MIGHT INFLUENCE mAb IMMUNOTHERAPY OF CANCER

Lorelei Irina Brasoveanu, Mihaela Daniela Simioana, Dan Hotnog, Viviana Roman and Mariana Claudia Constantin

Center of Immunology, Romanian Academy, Bucharest, Romania

Monoclonal antibodies (mAb) directed to tumor-associated antigens (TAA), which might activate human complement (C), started to be used in clinical trials with mAb immunotherapy, but the tumor cells were usually resistant to C-attack. Resistance to C-mediated lysis was associated to the presence of several C-regulatory proteins (CRP) such as CD59 (protectin), CD55 (DAF, decay-accelerating factor) and CD46 (MCP, membrane cofactor protein). Possible changes of antigenic expression might influence the role of CRP in C-lysis of tumor cells induced by anti-TAA treatment.

Therefore, the present study focused on evaluation of the antigenic expression of both therapeutical antigen targets associated with breast, ovarian and cervical cancer (MUC-1) or colon carcinoma (Ep-CAM) and MRC at the cellular membrane and molecular levels. The membrane expressions of TAA and MRC associated with human tumor cell lines derived from breast (SK-BR-3, MCF-7), ovarian (SK-OV-3), cervix (HeLA) and colon (COLO 201) adenocarcinomas were detected by flow cytometry in function of cell cycle phases of cytokine-treated tumor cells as compared to non-treated cells. The membrane expressions of TAA and CRP were correlated to mRNA expression, as detected by RT-PCR and using specific primers. In addition, we evaluated the role of differential levels of membrane, soluble forms and mRNAs of CD59 as compared to CD46 and CD55 in C-mediated cytotoxicity assays induced by anti-TAA mAb.

The levels of both proteic and genic expression of CRP might influence the effectiveness of immunotherapeutical strategies using mAbs directed to TAA: the effect could be increased if the resistance of cells is diminished or abolished. Knowledge of this data might help in selecting patients who would best benefit from the therapy and/or dimish tumor resistance to treatment, providing additional insights into the biological features of transformed cells that might impair the clinical efficacy of humoral immunotherapy.

**72**

## TOWARDS BIOCHEMICAL AND STRUCTURAL STUDIES OF THE HUMAN α7 ACETYLCHOLINE RECEPTOR EXPRESSED IN *PICHIA PASTORIS*

Vicky Katsemi, Konstantinos Poulas and Socrates J. Tzartos

Department of Biochemistry, Hellenic Pasteur Institute, Vas. Sofias 127, 11521 Athens;
University Department of Pharmacy, 26500 Patras, Greece

The human α7 nicotinic acetylcholine receptor (nAChR) belongs to the Cys-loop ligand-gated ion channel superfamily, which also includes serotonin, glycine, γ-aminobutyric acid type A and γ-aminobutyric acid type C receptors. It is a membrane glycoprotein and it has been shown to form homopentamers, that exert a cationic channel function. Each monomer contains an extracellular amino-terminal domain, followed by four transmembrane segments.

Neuronal AChRs play key roles in various neuron-neuron interactions and are, therefore, involved in many functions and serious neurological disorders, including Alzheimer's disease, Parkinson's disease, schizophrenia and others. In

order to therapeutically approach these disorders, the high resolution structure of the α7 AChR is required, although, the size and the hydrophobic character of the protein make it very difficult to structurally analyse. Cryo-electron microscopy studies on two-dimensional crystals of *Torpedo* AChR have given considerable structural information (1). Additionally, the high-resolution X-ray structure of the structural and functional homologue of the extracellular domain (ECD) of the AChR subunits, the acetylcholine binding protein (AChBP), is now available (2). AChBP forms a stable homopentamer and has 24% sequence similarity with human α7-ECD.

Several trials have been made to produce α7-ECD in different expression systems; however, the recombinant proteins, although soluble, were produced either in the form of microaggregates or in low amounts and were thus unsuitable for structural studies. The first successful production of glycosylated, soluble and ligand-binding human α7-ECD has been achieved in our laboratory, yielding 0.5 ± 0.1 mg purified protein per liter yeast culture (3). In case of the whole α7 nAChR, the human recombinant protein has only been expressed successfully in very low amounts.

In this study, first expression trials of whole α7 nAChR in *Pichia pastoris* have been carried out, resulting in the expression of the membrane protein, that can partially be solubilized. Moreover, a new construct of α7-ECD has been used in order to improve the quantity and quality of the expressed protein, thus enabling the production of high-amounts of homogeneous protein, that can be used for structural studies.

1 Miyazawa A, Fujiyoshi Y and Unwin N: Nature *423*: 949-955, 2003.

2 Brejc K, van Dijk WJ, Klaassen RV, Schuumans M, van der Oost J, Smit AB and Sixma TK: *411(6835)*: 269-276, 2001.

3 Avramopoulou V, Mamalaki A and Tzartos SJ: J Biol Chem *279(37)*: 38287-38293, 2004.

**73**

## FORMATION OF TRUNCATED PROTEINS AND HIGH-MOLECULAR-MASS AGGREGATES UPON SHORT (SOFT) ILLUMINATION OF CHLOROPHYLL BINDING PROTEINS

Paolo Antonioli[2], Sara Rinalducci[1], Natascia Campostrini[2], Pier Giorgio Righetti[2] and Lello Zolla[1]

[1]Department of Environmental Sciences, University of Tuscia, Viterbo;
[2]University of Verona, Department of Industrial Agricultural Biotechnologies, Strada Le Grazie 15, Verona 37134, Italy

Different spot profiles were observed in 2-D gel electrophoresis of thylakoid membranes performed either under complete darkness or by leaving the sample for a short time in low visible light, as occurs during laboratory manipulation. In this latter case, a large number of new spots with lower molecular masses, ranging between 15000 and 25000 Da, were observed, while high-molecular-mass aggregates, seen as a smearing in the upper part of the gel, appeared in the region around 300 kDa. Identification of protein(s) contained in these new spots by MS/MS revealed that most of them are simply truncated proteins derived from native ones, large fragments or high-mass aggregates.

Decreased intensity of staining with Coomassie blue, smearing and mobility shifts of protein bands in SDS-PAGE are phenomena usually observed when proteins are exposed to reactive oxygen species (ROS). This results from the formation of extremely reactive oxygen singlets generated by the exposure of chlorophyll-binding proteins of photosynthetic apparatus to low intensity light, conditions encountered during laboratory manipulation of samples prior to analysis. The oxidative process catalyzed by ROS can induce many different protein modifications, such as peroxide formation, side-chain product generation, backbone fragmentation in the $NH_2$ region and production of cross-links and molecular aggregates.

These data should alert researchers, who perform proteomic studies on complexes containing chlorophyll-binding proteins, to work in total darkness until complete removal of chlorophylls, and the possible presence of truncated proteins and other artefacts should be taken into account in data interpretation.

**74**

## PROTEOMIC ANALYSIS OF BLADDER CANCER CELL LINES FOR THE IDENTIFICATION OF MARKERS FOR METASTASIS

V. Prokos[1], K. Vougas[2], T. Katsorchis[1], J. Coleman[3], M. Fountoulakis[2] and A. Vlahou[2]

[1]Department of Biology-University of Athens;
[2]Department of Biotechnology, Foundation for Biomedical Research of the Academy of Athens, Athens, Greece;
[3]Urologic Oncology Branch, NCI/NIH, Bethesda, MD, U.S.A.

The management of bladder cancer could be significantly improved by the identification of protein biomarkers that mark the early onset of tumorigenesis as well as the acquisition of invasive and metastatic phenotype by the bladder cells. To address the latter, as a first step, we studied the proteome of the lineage-related T24 and T24M bladder cancer cell lines that differ in their metastatic potential. Total protein extracts from the cells were analyzed by two-dimensional gel electrophoresis. Image

analysis and spot quantification were performed by use of the PD-Quest software. Spot identification was carried out by MALDI-MS peptide fingerprinting. The protein map of these two cell lines generated through that process consists of approximately 400 different protein species. Multiple protein differences were observed between the two cell lines, which are currently being further evaluated for their potential involvement in the acquisition of the metastatic phenotype. Our future studies will also include investigation of the expression of these proteins in bladder cancer tissues as well as in biological fluids from bladder cancer patients.

**75**

**IDENTIFICATION OF BIOMARKERS IN BREAST CANCER BY PROTEOMIC ANALYSIS**

Ludovic Canelle[1,2], Cedric Pionneau[1], Jordane Bousquet[1], Raymonde Joubert-Caron[1] and Michel Caron[1]

[1]Laboratoire de Biochimie des Protéines et Protéomique (LBPP), EA 3408 Immuno-Pathologie et Immuno-Intervention, Université Paris13, UFR SMBH Léonard de Vinci, 93017 Bobigny, Cedex;
[2]R&D Immunoessais et Protéomique, bioMérieux, Chemin de l'Orme, 69280, Marcy l'Etoile, France

Breast cancer is a complex disease that still imposes a significant healthcare burden on women worldwide. The advent of advanced molecular biology techniques and the availability of proteomic strategies has opened up new opportunities and will potentially lead to the discovery of tumour biomarkers for early detection and prognosis of breast cancer. Pathological changes in breast cells might be reflected by proteomic patterns in serum. The identification in sera of tumour-associated antigens that elicit a humoral response is possible at the protein level; as well as autoantibodies biomarkers produced in response to the development of the tumour.

We screened antigens recognized by autoantibodies using two-dimensional (2-D) electrophoresis, 2-D Western blots with sera from cancer patients and from healthy people, and image analysis. Using optimised parameters to obtain reproducible protein patterns on 2D-blot, a set of 40 2D-blots was realized with 20 sera form patients with breast cancer (BCP) and 20 sera from healthy volunteers (HV). A matching experiment performed with image analysis software enabled the exact localisation of each relevant antigen hybridised by sera autoantibodies on the 2D-blots. Their position was recorded on preparative gels for subsequent identification by mass spectrometry (MS). Fifteen antigens were recognized by both BCP and HV sera and identified by MS. A total of seven antigens, which reacted preferentially with BPC sera, were identified by MS and MS/MS. Additionally, four antigens, detected with BPC sera, were identified as various isoforms of three proteins: G6PD (P11413), HSP70 (P08107) and DLDH (P09622). These results suggest that different post-translational modifications occur in tumour cells contributing to the humoral response observed against some specific protein isoforms, which could be associated with the cancer status.

**76**

**COMPARATIVE ANALYSIS OF SAMPLES OF COLORECTAL CANCERS: DEVELOPMENT OF AN APPROPRIATE STRATEGY**

Jordane Bousquet[1], Cedric Pionneau[1], Ludovic Canelle[1,2], Michel Caron[1] and Raymonde Joubert-Caron[1]

[1]Laboratoire de Biochimie des Protéines et Protéomique (LBPP), EA 3408 Immuno-Pathologie et Immuno-Intervention, Université Paris13, UFR SMBH Léonard de Vinci, 93017 Bobigny, Cedex;
[2]R&D Immunoessais et Protéomique, bioMérieux, Chemin de l'Orme, 69280, Marcy l'Etoile, France

The comparative analysis of two-dimensional electrophoresis gels is one of the main proteomic analysis steps. It makes it possible to reveal the presence or absence of some polypeptides linked to specific spot positions from complex images, to quantify the relative abundance of each of them and to control the validity of the differences using statistical tools.

We developed an original analytical strategy using the software Image Master 2-D Platinum (Amersham Bioscience, Orsay, France) in order to compare protein extracts from colorectal cancers of two ranks of malignancy (T2, T3). Differences in the relative abundance of polypeptides were searched across 2-D gels from cancers of T2 (N=4), or T3 (N=6) ranks. The number of spots was on average, 1000 ± 150, with percentages of matching from 55 to 65%, which is very satisfactory compared to the individual variability between patients. Coefficient of variation values (CV) were about 55% between the patients from the same rank, and approximately 80% between the two ranks. Thus, specific spots could be highlighted as features of a rank of malignance. Our strategy of analysis is fast and very efficient in comparing the polypeptide abundances between samples coming from different patients.

**77**

**STUDY OF NON-COVALENT INTERACTIONS BETWEEN ENDOGENOUS ANTIOXIDANTS AND AMYLOID-β-PEPTIDE BY MASS SPECTROMETRY**

Fotini N. Bazoti[1,2], Jonas Bergquist[3], Karin Markides[3] and Antonis Tsarbopoulos[1,2]

[1]Pharmacy Department, Laboratory of Instrumental Analysis, University of Patras, Patras 265 04;

[2]GAIA Research Center, The Goulandris Natural History Museum, 13 Levidou Street, Kifissia 14562, Greece;
[3]Department of Analytical Chemistry, Institute of Chemistry, Uppsala University, Box 531, SE-751 21, Uppsala, Sweden

Alzheimer's disease (AD), a leading cause of senile dementia, is associated with neurodegeneration, loss of cognitive ability and premature death. The cause of AD is presently unknown, but the existing hypotheses have centered on amyloid beta protein-containing senile plaques and protein tau-consisting neurofibrillary tangles, facilitated by a multitude of risk factors. Although, several proteins are involved in the pathology of AD, amyloid-β-peptide (Aβ), is currently believed to play a central role. Among the mechanisms proposed to justify Aβ's neurotoxicity, the prevailing hypothesis involves oxidative stress. Even if reactive oxygen species (ROS) are a consequence rather than a cause of Aβ aggregation, oxidative stress and free radicals can result in neurodegeneration and cell death. Therefore, antioxidants, particularly endogenous antioxidants such as Melatonin (M), which has been reported to possess neuroprotective and anti-amyloidogenic properties, may be beneficial in a therapeutic and/or preventive approach to AD.

Non-covalent interactions between proteins or peptides and ligands can reveal pathological conditions or be implicated in therapeutic approaches. Therefore, elucidating the structures and formation mechanism of these noncovalent complexes could be an important step towards understanding the disease processes, and the development of therapeutic approaches. In the present study, the possibility of non-covalent complex formation between Aβ and M was assessed by electrospray (ESI) mass spectrometric techniques (LC-TSQ-MS and LC-FTICR-MS). Mass spectrometry is a powerful tool for monitoring these interactions in real time, enabling the definition of the stoichiometry and the topology of the interacting species, which will assist in the development of a preventive or therapeutic regime for protection against the risk of AD. In addition, it will provide an effective screening method for testing the activity of possible anti-amyloidogenic indole analogs of M.

**78**

## COMBINING TEMPLATE MATCHING WITH SCALING INDEX AUTOMATED SEGMENTATION METHOD FOR MAPPING RIBOSOMES IN PROKARYOTIC CELLS BY CRYO-ELECTRON TOMOGRAPHY

A. Linaroudis, J. Ortiz, F. Förster, J. Kürner, C. Kofler, A. Briegel and W. Baumeister

Max Planck Institute for Biochemistry, 82152 Martinsried, Germany

The structure of the ribosome, a large RNA-protein complex, can be conveniently studied using X-ray crystallography, or cryo-electron microscopy in conjunction with single particle analysis. Although a very high resolution can be achieved with these techniques, information about ribosome localization and possible interactions with other proteins within the cellular context is lost, since complexes must be purified.

In order to localize 70S ribosomes within their physiological environment, cryo-electron tomography was applied to whole, ice-embedded cells from several organisms. In the case of *Spiroplasma melliferum*, separation of the membranes from the rest of the cell body was achieved using the scaling index method, which characterizes the dimensionality of the structural feature to which the voxel belongs. Appropriate thresholding of the calculated scaling indices creates a binary mask for each object group. Also, the ribosomes could be identified by the segmentation method, and their number was used as *a priori* information to drive the correlation process. Through template matching, which includes pattern recognition, 3-D alignment of identified particles and 3-D averaging, we established a map of ribosomes for each cell and, thereby, produced 3-D reconstructions of the corresponding ribosomes. Due to high contrast, parts of the membrane were identified as ribosomes through the correlation process, but, by having the membrane segmented, we could easily remove these false-positives.

In our 3-D reconstructions of *S. melliferum* (about 130 nm in diameter), ribosomes were clearly identified. Using 400 subvolumes, each containing one ribosome, a model was obtained at a resolution of about 4 nm, which showed several of the characteristic features of the subunits, *i.e.* the L7/L12 stalk, the central protuberance and the L1 stalk of the 50S subunit.

**79**

## BUILDING AN ONLINE ANTIBODY LOOP RESOURCE: A CDR RELATIONAL DATABASE, A CANONICAL CLASS PREDICTION SERVER AND AN ANTIBODY LOOP SEQUENCE/STRUCTURE REPOSITORY AVAILABLE TO THE STRUCTURAL AND COMPUTATIONAL BIOLOGIST

Dimitrios Nikoloudis[1], Sophia Kossida[2] and José W. Saldanha[3]

[1]Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, U.K.
[2]Division of Biotechnology, Center of Basic Research, Foundation of Biomedical Research of the Academy of Athens, Athens, Greece;
[3]Division of Mathematical Biology, National Institute for Medical Research, The Ridgeway, London NW7 1AA, U.K.

By varying the sequence and length of only six loops, antibodies can identify and bind virtually any antigen. These loops, called hypervariable or complementarity-determining regions (CDRs), were usually found to adopt one of a limited set of possible conformations, depending on the presence of a few key residues in the sequence. This observation gave birth to the canonical model, in which the three-dimensional conformation (or canonical class) of the corresponding loop could be predicted from sequence templates, with high accuracy for five of the six CDRs. Today, the increasing availability of new structural data presents an opportunity not only to improve the accuracy of the canonical model by identifying novel classes and reassessing the known ones, but also to study the basis of loop folding and gain insights into subtle antibody/antigen interactions. Steps are being taken in this direction that will enhance the capabilities of knowledge-based antibody engineering and assist attempts at *de novo* antibody designing. More specifically, an up-to-date structural repertoire of canonical classes is being acquired by clustering and analysis of all available antibody loop structures. An online dedicated database will provide antibody engineers, immunologists and structural biologists with a repository of categorised and cross-referenced loop conformations to be used as structural templates in modelling sessions, or for antibody/antigen interactions studies. The CDR repertoire will be automatically updated by new, publicly available antibody structures and will serve as a specialised sequence/structure resource open to computational biologists. Finally, the possibility to query the database with unknown Fv/Fab sequences will be implemented as a means to predict the conformation of CDRs from sequence alone, employing improved knowledge-based templates and sequence rules.

**80**

## AN ONLINE DICTIONARY SPECIALISED IN THE PROTEOMICS FIELD

K. Soderman[1], D. Barbouni[2], M. Fountoulakis[1] and S. Kossida[1]

[1]Division of Biotechnology, Center of Basic Research and [2]Library, Foundation of Biomedical Research of the Academy of Athens, Athens, Greece

The rapid growth of the proteomics field generates not only improvement of powerful technologies, mass spectrometry and chromatography, computational tools with methods to process, analyze and interpret copious amounts of data, but it also creates the necessity for a specialized vocabulary in order to deal with these fields. The need for translation (English to Greek/ Greek to English) of terms used in these areas is escalating rapidly.

The aim of this project was to create a comprehensive glossary of proteomics that will help the user to effectively integrate the scientific terminology and advance his/her concepts of this new scientific field. Our goals are to:

- Design a flexible, standards-based setting for development and publication of the glossary that allows continuous, internet-mediated contribution and peer review of content.
- Build a glossary from existing dictionaries and translated text books, with the interrelationships of references highlighted through hypertext links.
- Evaluate the glossary's content and format it to optimize its usefulness.

By engaging the expertise knowledge of proteomics and bioinformatics fields, we hope to create a tool which facilitates finding the correct word when needed.

**81**

## *IN SILICO* ANALYSIS OF EVOLUTION OF TUBULOINTERSTITIAL NEPHRITIS ANTIGEN MOTIFS

S. Kossida, V. Sideraki, C. Zervas, V. Kaltezioti, D. Beis and A. Charonis

Foundation of Biomedical Research of the Academy of Athens, Athens, Greece

Tubulointerstitial nephritis antigen (TIN-ag) is a basement membrane macromolecule expressed almost exclusively in the kidney and intestine. It was first identified in rabbit and mouse, followed by cloning of the human gene. Other TIN-ag-like proteins were later reported in *D. melanogaster* and *C. elegans*. A close relative to TIN-ag, TIN-ag-related protein (TIN-ag-RP), has recently been cloned and found to exhibit a wider tissue distribution, with high levels of expression in muscle tissue and placenta.

The aim of this study was to reconstruct the evolutionary history of the TIN-ag sequences. We undertook an *in silico* identification and analysis of TIN-ag and TIN-ag-related proteins across all known genomes, using information corresponding to putative structural and functional motifs present within the sequence. These were a furin cleavage motif, a disulfide isomerase/thioredoxin motif, a somatomedin-like motif and a nucleotide binding motif.

The findings suggest that there are only two related families of TIN-ag present in mammals. While the disulfide isomerase and somatomedin motifs are present in all members of both families during evolution, the nucleotide binding and furin cleavage motifs are only present in one of the two related families, namely the TIN-ag family.

These data suggest that TIN-ag may be more closely associated to the invertebrate TIN-ag gene, but no conclusive statement about the relatedness of the two

mammalian TIN-ag genes with the ancestral gene can be made. In order to gain further insight, experiments are under way to examine the TIN-ag homologue in *D. melanogaster*. These experiments may help us place the invertebrate gene closer to one family or the other, based on its expression pattern.

**82**
**AUTOMATICALLY ATTACHING FUNCTIONAL ANNOTATION TO HYPOTHETICAL PROTEINS IN ZEBRAFISH**

E. Zacharia[1], T. Karinos[2], A. Vlahou[3], D. Maroulis[1], M. Fountoulakis[3] and S. Kossida[3]

[1]Department of Informatics and Telecommunications, University of Athens, Athens;
[2]Department of Biology, University of Patras;
[3]Division of Biotechnology, Center of Basic Research, Foundation of Biomedical Research of the Academy of Athens, Athens, Greece

Identification of proteins after 2-D gel analysis and mass spectrometry is nowadays among the most commonly used procedures within the high-throughput proteomics field. Such an experimental procedure was used for the analysis of the zebrafish proteome. Data were analyzed and the currently available identification produced a few hypothetical proteins (HP). However, most of these HP are not really hypothetical.

Application of the blast algorithm and the use of both comparative genomics and motif detection approaches enabled us to attach functional annotation to most of the above-mentioned proteins.

A software package is described, showing the steps undertaken to resolve this issue automatically. A statistical analysis of the zebrafish results is also presented.

**83**
**COMPARISON OF THREE SOFTWARE PACKAGES FOR PROTEIN IDENTIFICATION OF MCF 7 MEMBRANE PROTEINS EXTRACTED WITH TRIFLUOROETHANOL**

Cédric Pionneau[1], Ludovic Canelle[1,2], Jordane Bousquet[1], Julie Hardouin[1], Michel Caron[1] and Raymonde Joubert-Caron[1]

[1]Laboratoire de Biochimie des Protéines et Protéomique (LBPP), EA 3408 Immuno-Pathologie et Immuno-Intervention, Université Paris 13, UFR SMBH Léonard de Vinci, 93017-Bobigny, Cedex;
[2]R&D Immunoessais et Protéomique, bioMérieux, Chemin de l'Orme, 69280 Marcy l'Etoile, France

The discovery of new markers or therapeutic targets for cancer implies the exploration of sub-proteomes to access proteins not visible by total proteome exploration. The membrane compartment is currently one of the most interesting. As they are often involved in cell homeostasis, membrane proteins represent targets for humoral immune response and tumor marker proteins in human malignancies.

Despite their biological importance, proteomic analysis of membrane proteins remains difficult. Recently, an elegant strategy based on the use of trifluoroethanol (TFE) organic co-solvent was applied to extract and separate *E. coli* membrane proteins. We adapted this strategy to investigate the membrane sub-proteome from the breast cancer cell line MCF 7.

The efficiency of this method was evaluated by SDS-PAGE, in-gel digestion and protein identification by nano LC-MS/MS. Sixty-eight distinct proteins from 13 gel plugs were first identified by Mascot (Matrix Science Ltd.). A classification of the proteins, based upon their respective SWISS-PROT annotations, showed that 53% of identified proteins are membrane or membrane-associated proteins, confirming the efficiency of the extraction method. In order to identify more proteins, two other software packages were used: Phenyx (Genebio SA) and SpectrumMill (Agilent). Phenyx takes maximum advantage of existing annotations in searched databases. With Phenyx, 11 new candidates not previously found with Mascot were identified (2 were membrane or membrane-associated proteins). Only 3 proteins were identified with Mascot but not with Phenyx. SpectrumMill is a suite of software tools with intelligent spectral extraction and interactive result validation. Using SpectrumMill, 13 new candidate proteins not previously found by Mascot were identified (3 were membrane or membrane-associated proteins), but 19 proteins were identified by Mascot but not by SpectrumMill. Considering a protein, the number of peptides identified can differ between different software packages. Then, sequence coverage was increased by merging the information provided by the different packages.

The TFE extraction method is an efficient way to extract membrane or membrane-associated proteins. The joint use of three identification software packages enables a more accurate and reliable identification.

**84**
**PROTEOME ANALYSIS OF LIVER FROM TRANGENIC MICE OVEREXPRESSING SMALL HETERODIMER PARTNER**

Kallia Kouyianou[1], Spiros Garbis[2], Ploumisti Dimitraki[2], Ioannis Talianidis[3], Michael Fountoulakis[2] and Georgios Tsiotis[1]

[1]Division of Biochemistry, Department of Chemistry, University of Crete, Knossos Avenue, P.O.Box 1470, GR-714 09 Heraklion;
[2]Foundation for Biomedical Research of the Academy of Athens, Soranou Ephesiou 4, GR 11527 Athens;
[3]Institute of Molecular Biology and Biotechnology FORTH, Vassilika Vouton, P.O.Box 1527, GR 711 10 Heraklion, Greece

The small heterodimer partner (SHP), an atypical orphan nuclear receptor, is a key regulator of genes involved in cholesterol-bile acid homeostasis and functions as a specific repressor of transcription. We studied differential protein expression in the liver of wild-type and transgenic mice expressing the human SHP gene. Proteins differentially-expressed in transgenic mice are thought to be affected by the higher levels of SHP. Liver protein extracts were analyzed by two-dimensional electrophoresis and the proteins were identified by MALDI-TOF-MS. Approximately 80 protein spots were differentially-expressed in the transgenic compared to the control mice. Major effects were evident in lipid accumulation, including the induction of a lipid receptor by a fatty acid-binding protein. Overexpression of SHP also triggered alterations in key enzymes of amino acid metabolism, nucleic acids, the urea cycle and was associated with changes in cellular stress proteins involved in calcium homeostasis, detoxification and proteins involved in protein repair and folding.

## Authors Index

(Figures refer to abstract number)