B.1 Comparative Expression Analysis between Proteomic and Transcriptomic Datasets Provides Novel Insights into TOR Signaling Regulation in Yeast

S. Bandhakavi¹, H. Xie¹, H. Sakurai², D.-H. Kim¹, and T. Griffin¹

¹Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN; ²School of Health Sciences, Faculty of Medicine, Kanazawa University, Kanazawa, Ishikawa, Japan

Although comparative expression analysis of biological pathways is frequently used in microarray experiments, it isn't always an option in quantitative proteomic profiling based studies because of the paucity of protein expression data obtained under various experimental conditions. In many instances, RNA and protein expression changes are not linear, but are often homodirectional (i.e., change in a similar direction). Thus a “qualitative” comparison of proteomic and transcriptomic datasets is likely to identify biological pathways and regulatory networks affected similarly under different conditions. As proof of principle, we have profiled protein abundance changes upon TOR inhibition by rapamycin treatment, and compared this data to existing expression information for corresponding gene products measured under a variety of conditions in Saccharomyces cerevisiae. Among the 127 proteins identified as showing abundance changes upon rapamycin treatment, a majority (112/127) demonstrated homodirectional transcriptomic changes under conditions of heat/oxidative stress. Because the known downstream responses regulated by TOR do not fully explain the extent of overlap between these two conditions, we hypothesized that activation of regulator/s of these stress responses cause or mimic TOR inhibition in yeast. Among the pathways tested, genetic analysis revealed that constitutive activation of heat shock factor 1 (HSF1) specifically inhibited TOR signaling and rapamycin resistance in yeast via elevated expression of unique target genes. Our results demonstrate the usefulness of comparative expression analysis between proteomic and transcriptomic datasets to unravel regulatory connections between biological networks and enhance the value of proteomic profiling studies.

B.2 Newly Configured Intact Protein-based Analysis System (IPAS) and Its Application in Plasma Proteomics Sample Handling

F. Xiang¹, H. Wang², W. Bankert¹, P. Winkler¹, and S. Hanash²

¹Shimadzu Biotech, Pleasanton, CA; ²Molecular Diagnostics, Fred Hutchinson Cancer Research Center, Public Health Sciences Division, Seattle, WA

Comprehensive profiling of the plasma proteomes has substantial relevance to the identification of circulating biomarkers that may have predictive value and utility for disease diagnosis. The substantial complexity and vast dynamic range of protein abundance in biological fluids, notably serum and plasma, present a formidable challenge. It turned out that sample processing was a major problem. Integration of multiple technologies is required to achieve high-resolution and high-sensitivity proteomics sample processing and analysis of biological fluids. We implemented an orthogonal multi-dimensional intact-protein analysis system (MD-IPAS), coupled with immunodepletion of abundant proteins and protein tagging, to quantitatively profile the human plasma proteome. Samples from carcinomas and matched controls were processed with newly configured MD-IPAS. Our results show that IPAS provides a combination of comprehensive profiling and quantitative analysis, with a substantial dynamic range, for disease-related applications. Applying IPAS technique enable to keep proteins intact during sample preparation and separation, which has important advantages compared with strategies that digest proteins at an earlier step. For example, post-translational modification (PTM) is believed to be involved in most (nearly all?) protein regulation and signaling, but digestion to peptides was obscuring the signature. Plus, the mass spectrometry was knocking off the weak bounded PTM groups. It was difficult to locate the sites of PTM, and if multiple sites were involved, one had an even harder time studying the distribution of the population of isoforms. The advantages of using IPAS include the ability to quantify and recover proteins as well as to assess post-translational modifications and cleavage. The data will also show the higher reproducibility, automation and versatility of the system.
The Trials and Tribulations of Producing Manuscripts Compliant with Journal Publication Guidelines

R. J. Chalkley and K. F. Medzihradszky
Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA

Developments in mass spectrometry methodology, instrumentation and software rapidly enabled in-depth characterization of complex mixtures, from protein identification, modification analysis to quantitation studies. The amounts of data produced in these studies changed the paradigm of data analysis to a point where manual verification of results is not generally practical. In addition, the user-friendliness of instrumentation and software made these technologies available to groups with no formal background in mass spectrometric instrumentation and analysis. These factors combined to cause a variety of problems within the proteomics community where studies were being published where the reliability of the experimental results could not be assessed independently.

Key representatives within the proteomics community met and a set of community agreed publication guidelines were formulated (1). The sponsoring journal “Molecular and Cellular Proteomics” is employing these guidelines for publication, but many authors are struggling to create manuscripts that are compliant with these instructions.

This presentation represents a tutorial discussing the guidelines in terms of why given pieces of information are required and provides examples of formats that would satisfy the requirements.

References

Identification of the Protein Biomarker Sequences of Non-Genomically Sequenced Bacteria by Composite Sequence Proteomic Analysis

C. Fagerquist
Western Regional Research Center, Agricultural Research Service, USDA, Albany, CA

High copy cytosolic proteins from bacterial cells can be used as biomarkers to classify genus, species, sub-species and strain when analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Amino acid variations within a protein biomarker across genus, species, sub-species and strains can result in variations in biomarker molecular weight (MW). The effect of these variations in a number biomarkers (20–50) can result in a unique MS profile or “fingerprint” of the microorganism. Data analysis of MALDI-TOF-MS spectra involves either pattern recognition or bioinformatic analysis. Bioinformatic analysis of MALDI-TOF spectra are increasingly popular due to the increasing number of bacterial genomes in publicly available databases. However, as the bioinformatic analysis relies exclusively on a comparison of genomically-derived protein MWs to MALDI-TOF-MS m/z peaks, the presence of post-translational modifications (PTM) can result in incorrect identification of a protein and lowered confidence in the identification of the pathogen. In addition, the previously mentioned variations in amino acid sequence can result in mis-identification of protein biomarkers if the bacterial strain has not been genomically sequenced. In consequence, it would be useful to identify the full amino acid sequence of a protein biomarker (including its PTMs) from non-genomically sequenced bacterial strains without the necessity of either genetically sequencing the biomarker gene or attempting full de novo MS/MS sequencing of the protein. A method is presented which demonstrates that it is possible to combine multiple proteomic identifications from genomically sequenced bacterial strains. The approach involves “bottom-up” proteomic identification of protein biomarkers of non-genomically sequenced Campylobacter strains by combining non-overlapping, sequence regions of proteomic identifications from genomically sequenced Campylobacter species/strains. It was possible to construct a “composite” amino acid sequence of a protein biomarker from non-genomically sequenced bacterial strains using multiple non-overlapping homologous sequence regions of phylogenetically proximate and distant species/strains. The new composite sequence was confirmed by both MS and MS/MS data.
Inactivation of Rat Liver Mitochondrial Aspartate Aminotransferase by a Halogenated Cysteine S-Conjugate.

M. Villar1, Z. Niatsetskaya2, A. Cooper2,3, and A. Artigues1

1Department of Biochemistry and Molecular Biology, University of Kansas Medical Center, Kansas City, KS; 2Burke Medical Research Institute, White Plains, NY; 3Department of Biochemistry and Neurology & Neuroscience, Weill Medical College of Cornell University, New York City, NY

Many electrophilic xenobiotics are detoxified through the mercapturate pathway by conversion to the corresponding glutathione S-conjugate, which, after sequential conversion to a cysteine S-conjugate and a N-acetyl S-conjugate, are subsequently excreted. However, a large number of haloalkanes produced on an industrial scale are nephrotoxic. This toxicity is due to the action of cysteine S-conjugate β-lyases. These enzymes irreversibly convert cysteine S-conjugates derived from haloalkanes to aminoacrylate and a reactive sulfur-containing fragment. Aminoacrylate rapidly undergoes tautomerization and hydrolysis to pyruvate and ammonia. The reactive sulfur-containing fragment [CF3HC(SiF)] derived from S-(1,1,2,2-tetrafluoroethyl)-L-cysteine (TFEC; the cysteine S-conjugate of tetrafluoroethylene) is known to thioacylate Lys residues of six kidney mitochondrial proteins in vivo including mitochondrial aspartate aminotransferase (mAAP). Rat liver mAAP catalyzes a β-lyase reaction with TFEC, resulting in irreversible, syn-catalytic inactivation of the enzyme. Fifty percent inactivation occurs after several thousand turnover events. Mass spectrometry analysis of the intact enzyme showed that inactivation of mAAP by TFEC results in multiple nonhomogeneous covalent modifications of the enzyme. In order to provide a detailed analysis of these modifications, trypptic peptide fingerprints were prepared and analyzed by tandem mass spectrometry on a LTQ FT mass spectrometer. Our results show that several Lys residues were modified by thioacylation resulting in a mass increase of 95.98 amu. In addition, one cysteine residue was noted to be modified by addition of 86.02 amu, consistent with Michael addition of aminoacrylate to a cysteine residue generating a lanthionine moiety. Finally, two Lys moieties in a single tryptic peptide were modified by addition of a fragment that resulted in a net gain of 99.99 amu for each lysyl moiety. This finding is consistent with the addition of a CF(HC(SiF)] to the ε-amino group of a Lys residue and elimination of H2S from CF3HC(SiF]. This type of addition has not been described previously. One of the modified residues (Lys 258) forms a Schiff base with PLP in the active site of the holoenzyme. Thus, the irreversible inactivation of mAAP by TFEC can be explained by the inability of the modified Lys 258 to form a Schiff base with PLP. All cysteine S-conjugate β-lyases found to date are PLP-dependent enzymes that are normally involved in amino acid metabolism. The toxicity of halogenated cysteine S-conjugates may be explained, at least in part, by inactivation of PLP-dependent enzymes and by covalent modification of proteins via reactions described herein.
Altered Proteome Biology of Cardiac Mitochondria Under Stress Conditions


1School of Medicine, University of California, Los Angeles, CA; 2European Molecular Biology Laboratory/European Bioinformatics Institute, Hinxton, United Kingdom

Myocardial ischemia-reperfusion (I/R) induces mitochondrial dysfunction and, depending upon the degree of mitochondrial injury, may lead to cardiac cell death. However, our ability to understand mitochondrial dysfunction has been hindered by an absence of molecular markers defining the various degrees of injury. To address this paucity of knowledge, we sought to characterize the impact of I/R damage on mitochondrial proteome biology. We hypothesized that I/R injury induces differential alterations in various mitochondrial sub-compartments, that these proteomic changes are specific to the severity of injury, and that they are important to subsequent cellular adaptations to myocardial ischemic injury. Accordingly, an in situ model of cardiac mitochondria injury was established to examine two stress conditions: reversible injury (induced by mild calcium overload, 100 μM, 5min at room temperature) and irreversible injury (induced by hypotonic stimuli, 5 mM Tris-HCl, 5 min at room temperature). Both forms of injury had a drastic impact on the proteome biology of cardiac mitochondria. Altered mitochondria function was concomitant with significant protein loss/shedding from the injured organelles. LC/MS/MS analysis identified 245 proteins from the mitochondrial milieu in the reversible injury group and 402 proteins from the irreversible injury group, among which 208 proteins were common to the released pool from both conditions. In the setting of calcium overload, mitochondria retained functionality despite the release of numerous proteins, and the majority of mitochondria remained intact. In contrast, hypotonic stimuli caused severe damage to mitochondrial structure and function, induced increased oxidative modification of mitochondrial proteins, and brought about detrimental changes to the sub proteomes of the inner mitochondrial membrane and matrix. Using an established in vivo model of regional myocardial I/R injury in mice, we validated key observations made by the in situ model. This pre-clinical investigation provides function and sub-organelle compartment information on a repertoire of cardiac mitochondrial proteins sensitive to I/R stress and highlights protein clusters potentially involved in mitochondrial dysfunction in the setting of ischemic injury.

Linking Proteolytic Activities to Mitochondrial Function; An Analysis of Cardiac Mitochondrial Degradome

J. Zhang, X. Li, J. N. Weiss, P. Korge, and P. Ping

School of Medicine, University of California, Los Angeles, CA

Proteases play pivotal roles in all living organisms. In mammals, proteases account up to 2% of human genome. In the yeast, proteolysis plays important roles in apoptosis, morphology, and protein turnover. Several investigations implicate that protease functions are coupled to their subcellular locations and are better defined in the context of the cellular milieu/local ancillary networks they reside. To date, organelle based distribution of degradome in mammalian cells are largely uncharacterized. Recently, the function of mitochondrial proteases has received more attention; at least several mitochondrial proteases were shown in the regulation of mitochondrial apoptosis. Given the profound importance of mitochondrial protease HTRA2. The other 50% of proteins were previously found only in non mitochondrial compartments such as cytoplasm, nucleus, and ER. To elucidate the impact of proteolysis on mitochondrial function, we tested whether proteolytic activities modulate mitochondrial function. Four protease inhibitors, bestatin hydrochloride, carboxypeptidase inhibitor (from potato tuber), chymostatin and iodoacetic acid, were examined. Among them, pretreatment with bestatin hydrochloride, a metalloprotease inhibitor selective for aminopeptidases, prevented calcium-induced mitochondrial swelling, indicating that proteolysis by this family of proteases may play a detrimental role in mitochondrial function. This study provides the first piece of evidence that cardiac mitochondrial function is coupled to proteolysis; it lays the foundation for future studies to explore protease-participated molecular mechanisms of mitochondrial dysfunction.
Characterization of p53 Post-translational Modifications in Cos Cells by Mass Spectrometry

A. Joubel¹, K. F. Medzihradszky¹, R. J. Chalkley¹, H. Hondermarck², and A. L. Burlingame¹

¹Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA; ²University of Sciences and Technologies, Lille, France

p53, a tumor suppressor protein, is involved in many signaling pathways and is the most frequently mutated protein in cancers. The activity of p53 protein is highly regulated through the interplay of many post-translational modifications, depending on the stimuli and stresses received by the cells. Cos cells contain a high level of p53, as it is stabilized due to its binding to the large-T SV40 antigen. Therefore they provide a suitable system for studying the post-translational modifications of p53. Utilizing CID analysis and high accuracy mass measurements a number of different modifications, both known and novel were deciphered. These included phosphorylation of serine and threonine residues, and acetylation and methylation of lysines.

This work was supported by NIH Grants NCRR RR001614, RR019934 and RR012961.

Integrated Systems Biology Study of Factors Influencing Amorphadiene Production from Metabolically Engineered Escherichia coli

F. Pingitore¹, C. Petzold¹, F. Nowroozi¹, L. Kizer², A. Redding¹, R. Dahl¹, W. Shui¹, D. Garcia¹, and J. Keasling¹

¹Department of Chemical Engineering, University of California, Berkeley, CA; ²Amyris Biotechnologies, Emeryville, CA

Introduction: Engineering exogenous biosynthetic pathways into microbes is an attractive, low-cost alternative to chemical synthesis or extraction from natural sources. However, the incorporation of the exogenous pathway often results in poor growth and significant metabolic changes in the host. Monitoring the microbe at the genetic, protein and metabolite levels reveals potential bottlenecks and perturbations that limit production of the desired metabolite. A combination of DNA microarray, proteomic, and metabolite approaches can unveil how Escherichia coli responds to a synthetic metabolic pathway in an effort to further our rational design of this organism. Our analysis of the amorphadiene-producing E. coli reveals broad effects associated with biochemical flux through the exogenous pathway impacting a wide range of host systems.

Methods: E. coli DH1 cultures were grown in M9 minimal media with 2% glucose as the sole carbon source. Protein samples were taken at three time points during growth (OD = 0.25, 0.6, 1.2). The cells were induced with IPTG immediately after the first sampling. The samples were centrifuged to pellet the cells and subsequently the supernatant was discarded prior to snap freezing the pellet with liquid nitrogen and placed in a -80 degree freezer until analyzed. For proteome analysis, the cells were lysed and the proteins were analyzed with a nanoLC-QTOF and TOF/TOF. Data analysis was accomplished with ProteinPilot and Mascot software packages.

Preliminary Data: More than 350 proteins have been identified from >7700 distinct peptides from E. coli expressing genes for the biosynthesis of amorphadiene. More than ten peptides accounting for ~50% sequence coverage are observed for each. Earlier work suggests that some of the terpene precursors may be toxic to the cells; consequently, a strain with an inactive pathway is used as a control. This also simplifies elucidating whether an observed stress is related to flux through the exogenous pathway or due to general stresses associated with protein overproduction. Preliminary data with both instruments clearly show that for the amorphadiene pathway an increased number of peptides, compared to the control strain, were detected.

Follow up experiments involving relative quantification with iTRAQ labeling of several time points during growth and comparison to DNA microarray data will be performed, so that investigation of pathway metabolites, relative protein abundances, and final product titers will be discussed.
Intelligent MS/MS Acquisition Results in Higher Sequence Coverage and More Confident Protein IDs from Complex Samples

J. Meza
Applied Biosystems, Foster City, CA

A major challenge in the study of proteomics is the analysis of complex samples. As sample complexity increases so does the need for a highly sensitive instrument with large dynamic range capabilities. Also required for deeper analysis of complex samples is the ability to perform high speed MS/MS scans without loss of spectral quality during an LC-MS/MS run. In the current study, Smart Exit (dynamic acquisition time of MS/MS spectra based on the evaluation of fragment ion intensity) and Dynamic Background Subtraction (in which precursor ion selection is based on the first derivative of intensity) were used in combination with a repetitive LC-MS/MS strategy that excludes previously acquired precursors in subsequent runs. Using this strategy, low attomole level sensitivity and four-log orders of dynamic range was achieved. Analysis of a yeast whole cell lysate resulted in the identification 739 nonredundant proteins with greater than 95% confidence from two 75 minute gradient LC-MS/MS runs. The use of dynamic acquisition time and the repetitive LC strategy combined with ProteinPilot software illustrate the effectiveness of our approach in identifying a large number of non-redundant proteins with high confidence in a relatively short time.

Separation of Stable Isotope-Labeled Unfolded Proteins Before MALDI MS Analysis Enables Quantification of a Wide Range of Serum Proteins

W.-L. Liao and I. Turko
Center for Advanced Research in Biotechnology, National Institute of Standards and Technology/University of Maryland Biotechnology Institute, Rockville, MD

Labeling of a large serum samples with stable isotopes in vitro can advance chances to recover and quantify low abundance proteins. However, this requires an affordable procedure of depletion and an affordable chemistry of modification.

In the present work, serum proteins were first fractionated by ammonium sulfate precipitation and then labeled with normal D₀-acrylamide and deuterated D₃-acrylamide. Fractionation with ammonium sulfate is a simple and well-reproducible procedure with a coefficient of variation lower than 10%. In addition, it can be performed side-by-side for a large sample size and for a large sample number. We found that a pellet of 30–40% saturation has only 7% of a total serum albumin whilst retains 60–100% of many other serum proteins. After labeling with acrylamides and mixing the D₀- and D₃-samples, the ratio for quantification is established, and it is safe to separate the proteins. However, isotope labeled proteins are unfolded, and a number of approaches to separate unfolded proteins is limited to various electrophoretic and chromatographic techniques in the presence of SDS and/or urea. We proposed a two-steps procedure that includes a preparative electrophoresis to separate proteins based on molecular mass with the following ion-exchange chromatography in the presence of 8 M urea to separate proteins based on charge. To underline the feasibility of this workflow to recover low abundance proteins, the 51 protein that have a concentration in serum lower than that of ceruloplasmin (0.3 mg/ml or 10⁻² to albumin) is listed. This includes many proteins secreted into extracellular space and several typical cytoplasmic proteins. We don’t know serum concentrations for most of these proteins and reasonably assume that the proposed workflow allows relative quantification of serum proteins at concentrations as low as 10⁻⁶ to albumin.

As a “proof of principle,” this robust workflow was used for the quantitative profiling of protein changes in serum associated with an exposure to room temperature for 24 hours compared to storage at 4 °C.
The 14-3-3 family of proteins is present in all eukaryotes, and bind a wide variety of protein partners. Due to these interactions, 14-3-3 proteins variously regulate the location, conformation and activity of their protein interaction partners. Homo- and heterodimers of 14-3-3 proteins bind their interaction partners through short linear motifs that usually contain a phosphoserine residue. Sets of 14-3-3 interacting proteins have been identified by affinity purification, yeast two-hybrid, and other methods. We have isolated the interaction partners of each human 14-3-3 isoform by tandem affinity purification (TAP) and identified sets of interacting proteins by shotgun proteomics.

Human 14-3-3 isoforms (beta, epsilon, eta, gamma, theta, sigma and zeta) were subcloned into a pCMV vector encoding an N-terminal TAP construct consisting of Protein A and calmodulin binding peptide separated by a tobacco etch virus protease cleavage site. HEK-293T cells were transiently transfected with each isoform and the 14-3-3 and associated binding proteins isolated by standard TAP methods. Isolated proteins were run on 1-D SDS-PAGE gels which were cut into bands, digested and subjected to LC-MS. MS data was searched using X!tandem and Sequest against the human IPI database.

A total of 1769 proteins were identified as binding to at least one 14-3-3 isoform, ranging from 313 interactors with 14-3-3 gamma to 724 with 14-3-3 beta. Analysis of gel slices containing native and recombinant 14-3-3 proteins indicates that while 14-3-3 sigma only forms homodimers, all other isoforms heterodimerise to some extent. 14-3-3 epsilon is the favoured heterodimerisation partner of the other isoforms. Hierarchical clustering of the interactors indicates that 14-3-3 sigma has the most distinctive pattern of protein partners, in accordance with its previously described distinct role in responding to DNA damage. Gene ontology analysis was used to identify characteristics of shared and isoform-specific interacting proteins. Characteristics of shared interacting proteins include regulation of apoptosis, protein folding and localisation. Distinct characteristics were also identified for interactors of individual 14-3-3 isoforms.

This dataset represents an important step in characterising the role of the functionally enigmatic 14-3-3 protein family. As a large set of physiologically relevant protein interactions, it will facilitate further studies of the dynamics of 14-3-3 binding; allow the identification of isoform-specific amino acid binding sequences in a cellular context; and help identify distinct roles for the multiple isoforms of this protein family.
B.15

Structural Analysis of Components of the Scorpion Venom *Vaejovis mexicanus smithi*

S. P. Salas-Castillo1, K. F. Medzihradzky1, C. F. Batista2, L. Possani2, S. Mahrus3, J. Wells3, and A. L. Burlingame4

1Department of Pharmaceutical Chemistry, University of California, San Francisco, CA; 2Biotechnology Institute, National Autonomous University of Mexico, Cuernavaca, México

Scorpion venoms constitute a vast source of bioactive peptides. A large number of toxins, mainly small peptides have been purified from the venom of different species and have been used as specific probes for neurotransmitter receptors and ion channels. The present study is focusing on the biochemical characterization of the venom of the Mexican scorpion *Vaejovis mexicanus smithi*, unique to the state of Morelos. This species has not been studied up to date.

2 mg of crude venom was fractionated by RP-HPLC and ~80 fractions were collected. Each fraction was subjected to MALDI-TOF mass spectrometry. The approximate size of different channel-blocking peptides has been reported. Fractions containing components of such molecular masses were selected for further studies.

Because of the complexity of the mixture and the size of the peptides of interest Edman sequencing as well as LC/MS, CID and ECD were utilized in order to decipher the amino acid sequence of the toxin components. We also present how a combination ECD analysis and selective N-terminal labeling can aid the de novo sequencing of large peptides.

This work was supported by NIH Grant NCRR RR001614.

B.16

Quantitative Proteomics Using Lanthanide Tags (MeCAT)

M. Linscheid1, R. Ahrends1, S. Pieper1, A. Kühn1, C. Scheler2, and H. Schlüter3

1Institut für Chemie, Humboldt-Universität zu Berlin, Germany; 2Proteome Factory AG, Berlin, Germany; 3Department of Endocrinology and Nephrology CBF, Charité-Universitätsmedizin, Berlin, Germany

We introduced the labeling approach based on lanthanide-complex tagging of proteins and peptides ("MeCAT"). The advantage over existing techniques is that this strategy allows the absolute quantitative determination of proteins and peptides very accurately in attomole quantities. The measurements are based on the precise determination of lanthanum amounts using combinations of LC/ICP-MS (for amounts) and ESI-MS/MS (for structures). Thus the way is paved to compare different proteomes not only by monitoring relative changes. It became feasible now to measure absolute quantities even for low abundant proteins and, since internal standards can be used, to compare different species, e.g., of cells, tissue compartments, bacteria and others.

Recently, we have used this scheme to determine changes in the proteome of eye lenses of pigs and different mouse types after 2D gel separations in analytical gels. The ultimate goal is to monitor changes causing crystallizations and cloudiness.

Since the superior detection capability of ICP-MS allows to segment single spots into small parts, e.g. to search for multiple proteins species in one spot—detailed analysis of otherwise concealed proteins in complex mixtures can be shown and “invisible spots” can be detected. To extend this approach, experiments are under way to use 2D-LC techniques combined with ICP MS and MS/MS detection in order to complement the gel separations used up to now.

In addition we have used the MeCAT approach to tag bioactive peptides. The labeled peptides serve to determine the changes in concentration in a quantitative fashion. By this means, enzyme activities can be found and accurately monitored at very low levels.
**B.18**

**Increasing the Depth to Which Complex Proteomes Can Be Penetrated by MuDPIT; Identifying Chromatographic, Mass Spectrometric, and Computational Improvements**


1St. Jude Children’s Research Hospital, Memphis, TN; 2Waters Corporation, Milford, MA

The depth of proteome penetration by MuDPIT is commonly believed to be limited by the overwhelmingly large number of different peptides in the mixtures subject to analysis. Yet standard multidimensional chromatographic procedures are seldom optimized to obtain the best possible results. We here report that major benefits can be gained by optimizing chromatographic resolution and sample loading in 2-dimensional peptide separations. 75-µm columns used for on-line mass analysis are easily overloaded, with significant reduction in the number of proteins identified per run. Although we observe that the quantity of protein that can optimally be committed to MuDPIT is thereby limited, we note that best scaling makes the technology suitable for the analysis of unexpectedly small numbers of cells. We also observe large increases in sequence coverage and numbers of proteins identified through optimal choice of fraction size in relation to chromatographic peak width in the 1st dimension of peptide separation. We utilize high capacity separations for both dimensions of peptide chromatography by employing stationary phases of 1.7 µm particle size. To exploit the high chromatographic resolution achieved, the rapid scanning capability of the Finnigan LTQ ion-trap mass spectrometer is used for on-line mass spectrometry. We describe improvements to the interface between UPLC and mass spectrometer that permit the strengths of both to be utilized. Results achieved with optimized 2-dimensional separations are presented for studies of 1) intrinsically unstructured proteins in mammalian cells, and 2) Class I-mediated antigen presentation in cells infected with influenza virus. Despite the improvements in peptide separation achieved, we observe that the number of assignments for peptide mixtures derived from these complex biological systems plateaus within the range of available peak capacities. Because peptide complexity remains high enough to overwhelm the real-time computational methods employed for dynamic exclusion in precursor ion selection, further improvements within the constraints imposed by available analysis time can be realized by expanding dynamic exclusion lists. Furthermore, by de novo sequencing we demonstrate that the incidence of false-positive peptide assignment is much higher than is generally appreciated, and provide an assessment of the factors responsible. Furthermore, the incidence of unassigned spectra is excessive. These observations point to the desirability of full disclosure of mass spectral data, and indicate that the need to develop accurate measures of the reliability of results obtained by MuDPIT persists.

**B.19**

**Proteome Analysis of Oxidative Stress Response to Cumene Hydroperoxide in Saccharomyces cerevisiae**

L. Tuli, A. Martins, W. Sha, P. Mendes, and V. Shulaev

Virginia Bioinformatics Institute at Virginia Tech, Blacksburg, VA

Recent advances in biomedical sciences and stress research have associated oxidative stress to several pathological conditions including cancer, ageing, autoimmune and neurodegenerative diseases and cardiovascular complications. As a common phenomenon to all aerobic organisms, oxidative stress has received intense scientific interest in the past few years. Described as a disturbed intracellular redox state, oxidative stress occurs when an imbalance transpires between existing reactive oxygen species and cellular defense mechanisms. Reactive oxygen species generated endogenously in response to a stimulus, react with biomolecules (DNA, proteins, lipids and carbohydrates) leading to cellular damage. Studies performed in model organisms reveal the intricate nature of a stress response, eliciting numerous biological processes each under a complex regulatory network. However, despite the ubiquity and importance of oxidative stress, we know little about the signals that trigger such a response and the temporal dynamics of the response itself. Saccharomyces cerevisiae has been a vital tool in stress research, with a wide application in fundamental and functional studies. We use yeast to study the kinetics of oxidative stress response to cumene hydroperoxide (CHP) from a systems biology perspective. We believe that information derived from yeast model can certainly be applied towards understanding stress response in higher eukaryotes.

Our integrative study monitored the stress response at three different levels: transcriptome, proteome and metabolome. In order to gain a dynamical insight on the stress response, samples were collected over a span of 3 to 20 min after addition of CHP. Changes in the intracellular redox environment, reflected by the modification of amino acids in proteins, allow proteins to act as stress sensors. Global protein changes were monitored using 2D (two-dimensional) gel electrophoresis in conjunction with LC-MS methodology. We used non-targeted protein profiling to focus upon proteins that changed significantly between treatments and those that depicted a time-treatment interaction effect.

Data obtained in this study shows that the dynamics of the oxidative stress response induced by CHP is a highly complex process and expands our understanding of plausible genes and proteins involved in this process. Clearly, the response to oxidative stress is dominated by two major events: a shutdown of growth and division, and induction of several processes involved with antioxidant protection.
Characterization of Cardiac Mitochondrial Subpopulations under Physiological Conditions following Free Flow Electrophoresis

O. Drews1, J. Zhang1, A.-S. Lee2, D. Liem1, X. Li1, and P. Ping1

1University of California-Los Angeles, School of Medicine, Los Angeles, CA; 2National Taiwan University, Taipei, Republic of China

Besides providing energy to the cells, mitochondria are involved in critical cellular events, including programmed cell death. Apoptosis is linked to the release of pro-apoptotic factors presumably by the mitochondrial permeability transition (MPT) pore. Studies on myocytes indicate intracellular subpopulations of mitochondria, which show different susceptibility to ischemia and pro-apoptotic stimuli. The mitochondrial heterogeneity seems to be linked to subcellular localization of the organelles in the intermyofibrillar and subsarcolemmal space. Classical isolation of these subpopulations requires a sequential extraction including protease degradation, which limits the assessment of their functional differences under physiological conditions, since mitochondrial surface proteins are degraded as well and impact their functional performance. Therefore, we isolated cardiac mitochondria by a single extraction method excluding enzymatic degradation, and subsequently employed zone electrophoresis on the Free Flow Electrophoresis system (ZE-FFE) to separate the mitochondrial subpopulations in parallel.

Previously, we have shown that cardiac mitochondria separated by ZE-FFE appear highly concentrated and purified in two major fractions [Mol Cell Proteomics (2006) 5, S21]. However, this separation was performed in a buffer optimized for resolution. When separated in a buffer optimized for assaying the susceptibility to MPT pore opening, a wide distribution of the mitochondria was obtained. Still, mitochondria in different fractions showed distinct responses in the assay. Here, we provide novel results on proteins associated with mitochondrial fractions and an optimized separation in the physiological buffer system, resulting in a similar separation as in the buffer optimized for resolution. All mitochondria separated in the physiological buffer system contained the inner mitochondrial membrane protein ANT and the outer mitochondrial membrane protein VDAC, indicating the mitochondria are intact and not stripped from their outer membrane. Furthermore, they showed full functionality in the MPT assay, which can be intervened by the MPT inhibitor Cyclosporin A. Differential proteomic analysis and identification by mass spectrometry of the proteins contained in the major mitochondrial fractions demonstrated association of sarcoplasmatic proteins with mitochondria, which were less deflected in the electrical field. Although this indicated the presence of the subsarcolemmal subpopulation, myosin heavy and light chain proteins were also associated with this fraction. In conclusion separation of mitochondria by ZE-FFE according to their charge states, as well as the proteomic analysis, demonstrate the existence of distinct mitochondrial subpopulations under physiological conditions.

Probabilities, Expectation Values, and Decoy Database Searching; Building a Unified Framework for Statistical Data Validation in Shotgun Proteomics

A. Nesvizhskii and H. Choi

Department of Pathology, University of Michigan School of Medicine, Ann Arbor, MI

Development of statistical methods for validation of peptide assignments to tandem mass (MS/MS) spectra obtained using database searching remains an important problem. In the past several years, several seemingly different approaches have emerged. The first approach is based on conversion of raw MS/MS database search scores into statistical measures called expectation values similar to those computed in sequence similarity search tools such as BLAST. Tandem, Mascot, and OMMSA are examples of tools performing such a conversion. The second approach involves searching composite databases containing decoy (e.g., reversed) peptide sequences. Decoy peptide matches are then used to estimate the global false discovery rate (FDR) and select filtering criteria that control for desired FDR. The third approach, exemplified by the computational tool PeptideProphet, combines multiple search scores to achieve improved discrimination, utilizes auxiliary (not database search related) information, and then constructs a Bayes classifier using probability mixture model.

In this work, we discuss the similarities, differences, and limitations of each method. In particular, we point out their complementarily and suggest a unified framework. In particular, we demonstrate that the probabilistic modeling approach of PeptideProphet and decoy strategy can be combined within a single semi-supervised framework, leading to improved robustness and higher accuracy of computed probabilities even in the case of most challenging datasets. The approach can be applied equally well to model the distributions of search scores before (raw scores) or after conversion to expectation values. Using several datasets of varying complexity, from control protein mixtures to a human plasma sample, and using three commonly used database search programs SEQUEST, MASCOT, and TANDEM/k-score, we illustrate how more accurate mixture estimation leads to an improved control of FDR in the classification of peptide assignments. The method not only allows estimation of global FDR, but allows calculation of posterior probabilities for each individual peptide assignment. Furthermore, it has a much higher discriminating power than the simple decoy database search strategy.

We also discuss the current limitation of PeptideProphet related to parametric assumptions in the model, and outline a full Bayesian approach that avoids restrictive model specification in the estimation of mixture distributions. Parametric specification of continuous mixture components is replaced by mixture of Gaussian mixtures that approximates the empirical score distributions well, leading to more accurate estimation of posterior probabilities. We show that the proposed method gives flexibility and robustness to the probabilistic validation in real data analysis, with a specific example of a high mass accuracy dataset generated using LTQ-FT instrument.
B.22
Automating an Integrated Protein Identification and Multiplexed Quantitation Workflow That Employs an Isobaric Mass Tagging Strategy
C. Parman, H. Javali, S. Lindberg, and B. Xie
PerkinElmer LifeSciences, Waltham, MA

A major barrier to quantitative analysis of complex protein mixtures is efficiently and reproducibly analyzing the enormous volumes of data generated. The necessary workflow must provide tools for the users to easily analyze large datasets in a consistent manner. To address the need, we have developed an automated analysis workflow for the identification and quantitation of proteins labeled with the ExacTag labels. These labels allow as many as ten different samples to be quantified simultaneously in a single mass spectrometry (MS) experiment. The workflow uses the Mascot search engine combined with proprietary quantification algorithms. By leveraging the automation capabilities of MatrixScience’s MASCOT Daemon, this process allows one to submit multiple data sets for unattended protein identification and quantitative analysis. This process is designed to be compatible with nearly every commercially available tandem MS instrument and most popular data formats.

We have applied these methods to several datasets of varying size and complexity. The process was validated using known protein mixtures. The results of these experiments show average CVs between replicate measurements to be less than 20% and quantification to within 10% of the known input ratios. In addition we show results of replicate measurement on two MS platforms agreed with each other to within 5%.

The capabilities of the approach for the analysis of large datasets are demonstrated by analysis of time course data from the Hela cell nucleolar proteome after treatment with the metabolic inhibitor actinomycin C. Cells were collected over seven time points and labeled with the ExacTag reagents. The resulting dataset consisted of over 50 MS/MS runs from 27 gel slices. Both the protein identification and quantitation results are presented.

B.23
Quantitative Tissue Proteomics and Biofluid Metabolomics of Cholangiocarcinoma
Y. Wang1, A. Sharif2, S. Taylor-Robinson2, and W. Griffiths1
1University of London, The School of Pharmacy, London, United Kingdom; 2Faculty of Medicine, Imperial College, London, United Kingdom

Cholangiocytes are the epithelial cells which line the bile ducts. Cholangiocarcinoma (CCA) now accounts for 3% of all gastrointestinal cancers and is the second commonest primary hepatic tumour worldwide. The incidence of intrahepatic cholangiocarcinoma (ICC) has increased dramatically in the UK. This tumour has a very poor prognosis and the aetiology is not clearly understood. The condition usually presents late and is advanced at the time of diagnosis, making it unresectable. It therefore has a high mortality and poor prognosis. Early markers of the disease are lacking and are needed to diagnose this cancer early, and hence improve overall survival. Oxysterols, bile acids, and environmental toxins, metabolised by the phase I and phase II drug-metabolising enzymes expressed in the hepatocyte and cholangiocyte, have been implicated in cholangiocarcinogenesis. Bile acids enhance cholangiocyte proliferation and alterations in biliary content may serve as disease biomarkers. The link between bile acids and CCA has encouraged us to employ both proteomic (to profile protein changes in the cholangiocyte) and lipidomic (to profile the bile acid content of bile) approaches in the study of CCA.

In our experiments we have compared the proteomic profile of periduc- tal infiltrating ICC and normal adjacent liver. We used the iTRAQ peptide tagging approach. Of particular interest is the observation of altered expression of oxysterol binding proteins (OBPs) in ICC tissue. OBPs have been implicated in the regulation of cholesterol synthesis via Sterol Regulatory Element Binding Proteins (SREBPs). Down regulation of OBP could lead to excessive cholesterol synthesis, and bile acid formation. Bile acids activate the epidermal growth factor receptor (EGFR) and EGFR activation leads to disordered cell cycling of cholangiocytes, ultimately leading to cholangiocarcinoma. Of further interest is the down regulation of many CYP enzymes and UDP-glucuronosyl transferases, but up-regulation of bile salt sulfotransferase and UDP-N-Acetyl glucosamine transferase. Significantly, the levels of sulphated bile acids are elevated in bile from CCA patients.
B.24

Subtle Modification Isotope Ratio Proteomics; Manipulation of $^{15}N/^{14}N$ ratio in Whole Plants Using Hydroponics

A. Laganowsky¹, J. Nishio², S. Bassilian¹, P. Souda¹, J. Katz³, K. Faull¹, and J. Whitelegge¹

¹University of California Los Angeles, Semel Institute, Los Angeles, CA; ²College of Natural Sciences, California State University, Chico, CA; ³Spielberg Family Center for Applied Proteomics, Cedars-Sinai Medical Center, Los Angeles, CA

Subtle modification of isotope ratio proteomics (SMIRP) has been proposed as a minimally intrusive in vivo metabolic labeling strategy for expression proteomics, to ultimately include living humans. Carbon or nitrogen stable isotope ratios are altered minimally such that typical peptide isotopomer distributions are significantly modified, yet not so much as to dramatically extend the envelope. The isotope ratio is calculated from the intensities of individual isotopomer peaks, and knowledge of the peptide atomic formula. Where two samples have been mixed for relative expression measurements, the measured composite isotopomer distribution is deconvoluted to yield the relative contribution of each sample. In the case of protein turnover measurements, the contribution of old versus newly synthesized protein is deconvoluted. Importantly, the SMIRP strategy relies upon a single isotopomer distribution, keeping mass spectra as simple as possible with benefits for sensitivity and ensuring that all tandem mass spectra are useful. The model plant Arabidopsis thaliana was grown in a hydroponic system that allowed complete mineral nutrition manipulation. SMIRP was used to evaluate the plant proteome under iron-sufficient and iron-deficient conditions. Here manipulation of the $^{15}N/^{14}N$ ratio was achieved by the addition of 3% atomic $^{15}N$-nitrate to the hydroponic growth medium. After incorporation bias, the final increase in plant $^{15}N$ content was close to 2%, a value very close to that we believe to be the optimum isotope ratio change for $^{15}N$ experiments. Unlabeled and labeled iron-sufficient versus iron-deficient thylakoid membrane proteins were analyzed by combined liquid chromatography-mass spectrometry and tandem mass spectrometry to evaluate the effects on both intact protein mass and more importantly peptide isotopomer distributions. This presentation will focus on the deconvolution of mixed isotopomer distributions to yield relative expression information as the plant thylakoid responds to iron-deficiency stress.

B.25

Identification of Proteolytic Processing Sites in Human Serum by Specific, Enzymatic Labeling of N-Termini

P. Wildes and J. A. Wells

Department of Pharmaceutical Chemistry, University of California, San Francisco, CA

Proteolytic processing is a key post-translational modification in blood plasma and serum. Most secreted proteins are processed by proteases at least once, when the signal peptide is removed. Additional processing occurs in many cases: when, for example, a single precursor is processed into multiple chains or when extracellular domains of membrane proteins are shed. Finally, proteolysis is key to the activation and propagation of multiple extracellular signaling cascades, including blood coagulation and complement activation. Despite its importance, progress in identifying proteolytic processing sites on a proteomic scale has lagged behind identification of other post-translational modifications, due to the lack of a good chemical “handle” for positive enrichment of processed peptides. Here we present a method to specifically label free N-termini created by proteolytic processing using subtiligase, an engineered enzyme with peptide ligase activity. Human serum or plasma is enzymatically labeled with biotin, digested with trypsin, and N-terminal peptides are captured on streptavidin beads. The peptides are released by digestion with TEV protease, leaving a two-residue “signature sequence” on the N-terminus aiding identification by LC/MS/MS. In preliminary experiments we have identified many annotated N-termini in, for example, coagulation and complement factors, as well as proteolytic processing sites that are not annotated in genome databases. We expect this work to lead to a more complete map of proteolytic processing sites in blood plasma and serum. In addition, this method is compatible with isotopic quantitation methods like iTRAQ, and will be useful for identifying proteolytic cleavages in serum that may be biomarkers of disease states.

B.26

Profiling Protein Expression in Primary Tissues Using New Informatics Tools for Label-Free Quantitative Proteomics

P. Cutillas and B. Vanhaesebroeck

Centre for Cell Signalling, Queen Mary University of London, United Kingdom

Label-free quantitative mass spectrometry presents attractive features as a tool for functional proteomics, including (i) the ability to compare an unlimited number of samples, (ii) the applicability to primary tissues and cultured cells, (iii) its straightforward workflow without chemical reaction steps, and (iv) its usefulness not only for relative quantification but also for estimation of absolute protein abundance. However, implementation of label-free quantitative proteomics requires of robust software for the automation of data analysis, which is not widely available. We developed a computer program that reports ion intensities of all the peptides identified in data-base searches and their comparison across samples. Thus in contrast to the ‘spectral-count’ approach sometimes used to derive semi-quantitative information, the approach used here also considers the ion intensity value of each of the identified peptides, with the consequent increase in accuracy of the quantitative information. By comparing ion intensity values of the identified peptides across samples, the approach shows levels of precision comparable to those afforded by label-based methods. In order to test the performance of the created software and analytical strategy in a challenging model, we applied this approach to quantitatively characterize the most abundant proteins in murine brain, heart, kidney, liver and lung. We matched 8,800 MS/MS peptide spectra to 1,500 proteins and generated 44,000 independent quantitative data points to profile the ~1,000 most abundant proteins in mouse tissues. This dataset provides a quantitative profile of the fundamental proteome of a mouse, identifies the major similarities and differences between organ-specific proteomes, and serves as a paradigm of how label-free quantitative MS can be used to characterize the phenotype of mammalian primary tissues at the molecular level.
B.27

Direct Measurement of Fluorogenic Substrate in Crude Enzyme Membrane by LC/MS

F. Gharahdaghi, M. Stein, M. Bock, C. Scott, and J. Arriza
AstraZeneca, Wilmington, DE

A homogeneous high throughput assay was developed for identification of enzyme inhibitors using a substrate, decanoyl 7-amino-4-methyl coumarin (D-AMC). The substrate is cleaved by the enzyme to release the fluorescent molecule 7-amino-4-methyl coumarin (AMC), providing a direct measure of enzyme activity. However, this fluorophore has relatively short wavelengths for maximal excitation and emission (355 nm and 460 nm, respectively), which can be problematic when screening compound collections due to quenching (false positive) and intrinsic fluorescence (false negative) artifacts of the test compounds. A separation-based liquid chromatography/mass spectrometry (LC/MS) approach has been investigated as an alternate detection method that does not have these liabilities. A sensitive, specific and robust LC/MS method was developed for quantitative measurement of AMC. We will present the method validation, including selectivity, calibration model, accuracy, and extraction efficiency. This assay was further validated using known inhibitors and EC50 values agreed with results using the homogeneous fluorescence assay. Compounds that were incompatible with the fluorescence assay did not cause interference in the LC/MS method. The LC/MS assay has been converted to a 96-well plate format using automated liquid handling systems to enable higher throughput screening of test compounds.

B.28

Phosphorylation-specific MS/MS Scoring

S. Payne1, M. Yau1, M. Smolka2, H. Zhou2, and V. Bafna1
1Department of Computer Science Engineering, University of California, San Diego, La Jolla, CA; 2Laboratory of Proteomic Biology, Ludwig Institute for Cancer Research, San Diego Branch, La Jolla, CA

Tandem mass spectrometry is a powerful tool in the discovery of protein phosphorylation sites, and has been widely used to annotate the phosphoproteome of both whole cells and sub-cellular fractions. Facilitating these studies have been experimental advancements that enrich phosphorylated peptides from complex mixtures, without which phosphorylated peptides would constitute less than 1% of the biological sample. Here we present a new scoring function for the InsPecT software, rigorously trained to understand and score phosphorylated peptide tandem mass spectra, without the need for manual validation.

Protein mass spectrometrists have long recognized the characteristic shape or profile of several ion types. Common examples include: proline directed fragmentation; the correlation between a y-h2o and peaks; y peaks intensity related to spectrum position. All of these ideas (intensity value, spectrum region, flanking residue, related peak presence) can be abstracted and encoded into an ion profile containing the probabilities of fragment ion generation. We have chosen Bayesian networks to represent the probabilities inherent in ion fragmentation. The relationships between ion fragments intuitively relates to a Bayesian-style prior probability model, storing the probability of an event based on other related events, e.g., P(y-h2o | y). The addition of peak scoring dramatically improves the discriminatory ability of the scoring function. Additionally we have written a new parent mass correction routine for the program, a crucial requirement for tag-base searching.

We compared our new algorithm against Comet and Xtandem on a test dataset of 6400 MS/MS spectra from yeast S. cerevisiae enriched using IMAC. The run time on a desktop PC was: 30 minutes for InsPecT, 8 hours for Xtandem, and 33 hours for Comet. Results were filtered to 5% false positive by comparison to a decoy database. 385 phosphopeptide spectra were identified by all three programs, and ~900 additional spectra were identified by at least one program. Inspect identified the greatest total number of phosphorylated spectra. Additionally, the spectral annotations not identified by InsPecT tended to score poorly on the trained scoring function, and did not contain characteristic features of phosphopeptides when manually inspected.