High-Performance Mass Spectrometry Facility

The High-Performance Mass Spectrometry Facility (HPMSF) provides state-of-the-art mass spectrometry (MS) and separations instrumentation that has been refined for leading-edge analysis of biological problems with a primary emphasis on proteomics. Challenging research in proteomics, cell signaling, cellular molecular machines, and high-molecular weight systems receive the highest priority for access to the facility. Current research activities in the HPMSF include proteomic analyses of whole cell lysates, analyses of organic macro-molecules and protein complexes, quantification using isotopically labeled growth media, targeted proteomics analyses of subcellular fractions, and nucleic acid analysis of ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) oligomers. More than a dozen microbial systems are currently being studied in the HPMSF by researchers from throughout the country. In addition, there are several ongoing projects in higher order systems (including mammalian systems) that are investigating a broad range of biological questions from cancer screening to infectious diseases to fundamental questions of post-translational modifications and protein-protein interactions.

Capabilities

The facility provides a complete suite of mass spectrometers for proteomics analysis. The available instruments range from a group of five ion trap spectrometers for tandem mass spectrometry (MS/MS) work to very high-sensitivity and high-resolution Fourier transform ion cyclotron resonance (FTICR) spectrometers that offer 1 ppm mass measurement accuracy. Proteomics analysis on these FTICR spectrometers is complemented by a quadrupole time-of-flight (QTOF) spectrometer that combines MS/MS analysis with mass measurement accuracy. These spectrometers are coupled with very high-resolution separations (500 peak capacity) that are highly beneficial to these areas of research. Focused research projects into biomolecular complexes and macro-molecules are supported on the facility's 7-T FTICR spectrometer, a flexible instrument that can be configured in many different ways.

Instrumentation & Capabilities

Mass Spectrometers

- Quadrupole time-of-flight
- Five ion trap spectrometers
- Three linear ion trap mass spectrometers

FTICR

- 11.5-T FTICR
- 7-T FTICR
- 9.4-T FTICR
- 3.5-T FTICR

Additional Capabilities

- Ten custom high-performance liquid chromatography (HPLC) systems
- Agilent capillary HPLC system

The HPMSF is committed to maintaining state-of-the-art MS and separations capabilities. To this end, the facility's staff work to develop new capabilities such as the ion funnel, dynamic range enhancements applied to MS (DREAMS), and data analysis tools, which are incorporated into the capabilities of the facility as they become available. This work is supported by a separate 3.5-tesla FTICR spectrometer, so

that the work of sample analysis is not interrupted by this developmental work. The members of the facility staff are highly skilled in all areas required for proteomics research, from sample preparation to analysis and data interpretation, and they are available to help develop methodologies to tackle these challenging problems. As needed, scientific staff not assigned to the facility can be accessed as matrixed members of the facility. Since its inception, more than 100 separate user projects have been undertaken in the facility, with some spanning over a year in duration.

The HPMSF has developed state-of-the-art software for the acquisition and analysis of FTICR mass spectra. This software package is called ICR-2LS. It is a Microsoft Windows-based application that enables many of the unique instrument control functions developed in the W.R. Wiley Environmental Molecular Sciences Laboratory (EMSL). This same software allows automated spectral interpretation of raw FTICR data. The spectral interpretation features are integrated into our application and, thus, are not easily transferred. The data acquired in the facility is stored and managed in an in-house-developed relational database.

Mass Spectrometry Research Capabilities. Five ion trap mass spectrometers from ThermoElectron are available in the facility: two Finnigan LCQ Classics, two Finnigan LCQ Duos that provide improved sensitivity, and one Finnigan LCQ DECAXP that provides the latest in sensitivity improvements from Finnigan. The ion trap instrument is a three-dimensional quadrupole ion-trap-based instrument designed for use with electrospray ionization sources. These instruments are well-suited to MS/MS spectrometer experiments because of their very high collection efficiency for product ions. The mass range of this instrument is 150 to 2000 m/z, but the range can be extended to 4000 m/z for some applications. The LCQ instrument has a maximum resolution of 10,000 in the zoom-scan mode, and 4000 in full-scan mode. In addition, the system is easily operated in either positive or negative ion mode with the addition of SF_6 as a sheath gas.

Figure 1. Ion trap mass spectrometers.

Fourier Transform Ion Cyclotron Resonance Mass Spectrometers. This ultra-high performance mass spectrometer uses a wide-bore (205 mm), passively shielded 11.5-T (Figure 2) superconducting magnet. The spectrometer is equipped with an electrospray ionization source and an ion funnel. The 11.5-T FTICR has a resolution of 150,000 at m/z 60,000 and a mass accuracy of 1 ppm for peptide samples with molecular weights ranging from 500 to 2000 Da. Ions are collected external to the magnet in a series of quadrupoles that allow the researcher to eliminate irrelevant ions before analysis in the FTICR. DREAMS is a unique capability of this facility. The 11.5-T FTICR instrument can be fitted with an HPLC system and is equipped with an infrared laser for multiphoton dissociation of samples for MS/MS in the ion cyclotron resonance (ICR) cell.

Figure 2. Wide-bore, passively shielded 11.5-T FTICR mass spectrometer.

The 9.4-T FTICR spectrometer (Figure 3) is a 150-mm bore, actively shielded Bruker Daltonics APEX III. The original ion source has been replaced with a custom source that incorporates a dual-channel ion funnel for simultaneous introduction of calibrant ions, DREAMS technology, and automatic gain control. These modifications permit maximum use of the spectrometer's capabilities by maintaining the optimum number of ions in the ICR cell throughout a liquid chromatographic (LC) separation. The outstanding resolution of over 60,000 is maintained throughout the separation with a sensitivity comparable to the other FTICR instruments in the facility. This system is seamlessly integrated with the facility's automated HPLC (5000 psi) system for unattended operation 24 hours a day, 7 days a week.

Figure 3. 150-mm bore, actively shielded 9.4-tesla FTICR instrument.

The Waters' Micromass QTOF Ultima API is an orthogonal extraction-TOF mass spectrometer that has enabled automated exact mass measurement with the ultimate performance (Figure 4) in MS and MS/MS. The instrument features optimized resolution, enhanced ion optics for enhanced sensitivity, a quadrupole mass filter, and a collision cell for MS/MS analyses. These features are also accessible over an enhanced linear dynamic range, enabling rapid and reliable quantification. This powerful combination of capabilities delivers simple exact mass measurement of precursor and fragment ions with maximum sensitivity to yield the highest confidence in structural elucidation and databank search results.

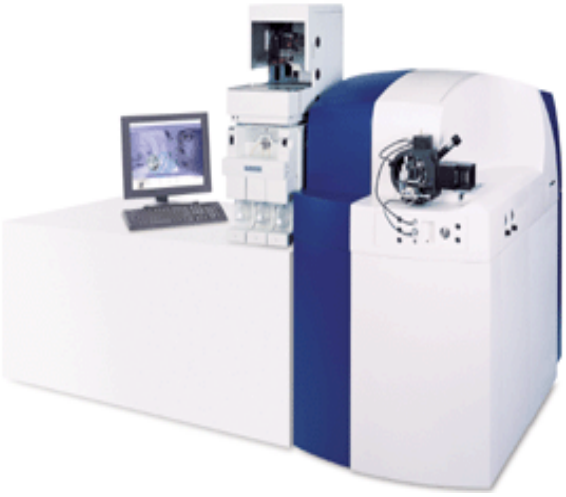


Figure 4. QTOF instrument.

A signature capability of the facility is the efficient coupling of capillary separations (Figure 5) to our mass spectrometers. Instruments for both LC and capillary electrophoretic (CE) separations are available. Two ABI 270A electrophoresis systems are available along with one Agilent capillary LC and one Shimadzu capillary LC system. Unique to this facility are the in-house-developed LC systems. These systems deliver constant-pressure gradient separations at up to 5000 psi. Continuing developments of high-pressure versions of these LC systems feature a PAL autosampler with cooled sample holder, VALCO high-pressure valves, and ISCO syringe pumps. Computer software has been developed that allows the system to be configured with any of our spectrometers through DCOM communication protocols.

Figure 5. Capillary ultra-high pressure separations capability.

The 7-T FTICR mass spectrometer is based on a 160-mm-bore, superconducting magnet and is equipped with a custom ESI source (Figure 6). This instrument has high mass resolving power (e.g., a mass resolution greater than 2,000,000 has been obtained for insulin), while unit resolution is routinely achievable during online capillary isoelectric focusing (CIEF) experiments for proteins with mass resolution less than 30,000 μ; mass accuracy less than 5 ppm is typical for peptide/protein samples with molecular masses ranging from 500 to 30,000 μ. A detection limit of approximately 10 attomoles has been obtained with online LC and CIEF separations.

Figure 6. 7-T FTICR mass spectrometer.

The high-performance FTICR mass spectrometer is coupled to a 3.5-T, 330-mm-bore magnet and is equipped with an electrospray ionization source. The 3.5-T FTICR instrument (Figure 7) has a mass resolution of 50,000 to 150,000 and a mass accuracy of 3 to 7 ppm for protein samples with molecular weights ranging from 5000 to 20,000 Da. The ion optics on this system are the most highly developed in the HPMSF. The spectrometer incorporates DREAMS technology and a dual-channel ion funnel for simultaneous introduction of a calibrant. The detection limit of the instrument has recently been improved down to 3 x 10^{-20} moles or about 18,000 molecules (based on sample consumption). The year 2004 was the last year of this instrument's use. It was replaced by an FTICR spectrometer with a 12-T magnet, intended for top-down proteomics (see Upgrades).

Figure 7. 3.5-T FTICR mass spectrometer.

Upgrades

Several important enhancements of HPMSF capabilities were made in 2004.

A new BRUKER DALTONICS APEX III data station was installed on the 11.5-T FTICR spectrometer. This upgrade permits automated operation of the spectrometer with our custom HPLC platforms.

Figure 8. 12-T mass spectrometer.

The acquisition and installation of a 12-T, 110-mm-bore actively shielded magnet was completed in the fall of 2004 (Figure 8). This magnet replaces the 3.5-T FTICR spectrometer in the facility. During the coming year, the vacuum system for the magnet will be installed and set up for the analysis of intact proteins by electron capture dissociation. This top-down analysis will complement the bottom-up analysis that the facility has offered, and will provide a more complete proteomics analysis capability. The top-down analysis helps researchers understand the role, function, structure, and changing nature of complex cellular proteins.

Cellular Response to Protein Nitration in C2C12 Myocytes Detected with Complementary Approaches: Cell Biology and Global Proteomic Screens

DJ Bigelow,[(a)] CA Sacksteder,[(a)] DL Stenoien,[(a)] TV Knyushko,[(a)] MP London,[(a)] B Gopalan,[(a)] DG Camp,[(a)] and RD Smith[(a)]

(a) Pacific Northwest National Laboratory, Richland, Washington

Studying individual proteins has reached its insightful limits. Progress now requires the study of systems of proteins and how they interact to affect a biological change. Having a more complete picture of all the proteins that participate in a particular change in a cell is critical to understanding how a system of proteins interact. This study identifies many new proteins that are involved in the addition of nitrogen to proteins for both normal and abnormal cell behavior.

We have identified 3-nitrotyrosine-modified proteins from gel-free global proteome analyses of skeletal muscle myocytes in culture. This proteomic screen is a critical step toward understanding cellular dysfunction when nitrated proteins accumulate in heart and skeletal muscle through aging, cardiovascular disease, and environmentally induced inflammation. Nitration of proteins by 3-nitrotyrosine modification signals the presence of cellular peroxynitrite, a potent and reactive oxygen species. Increased levels of nitration have been observed in more than 80 different pathologies, implicating nitrotyrosine modification in cellular damage. Not all nitration is necessarily pathogenic, and recent observations of reversible nitration-denitration of metabolic enzymes suggest a role for nitrotyrosine modification in redox-sensitive regulatory processes. Determining the extent to which nitration occurs in normal cells will be an important requirement if we are to understand when nitration leads to pathology in cells and when it is part of normal regulatory processes.

As a step in this direction, we recently detected the ability of myocytes to undergo a rapid loss of protein nitration after exposure to high doses of peroxynitrite (Figure 1). This denitration is 75% complete within two hours, followed by a slower phase of recovery to basal levels. Treated myocytes remain attached to the substrate and are fully viable up to 72 hours after treatment, suggesting that the observed denitration is a normal survival response to nitration rather than a pathological response leading to cell death. Lactacystin, an inhibitor of the proteasome, inhibits only about one-third of this denitration, indicating that multiple cellular mechanisms, including proteasome-dependent pathways, are involved. Current efforts are aimed at defining these additional processes.

This observation of an immediate cellular response to nitration of proteins has not been previously reported, but it likely represents mechanisms that are critical for the ability of heart and skeletal muscle to maintain protein function under the highly aerobic metabolic conditions accompanying contractile activity. In all likelihood, these denitration processes represent a second line of defense that operates under conditions of high oxidative stress after the first line, the antioxidant defenses, are overwhelmed. Elucidating the molecular mechanisms of denitration will provide insights into the steady-state accumulation of nitrated proteins in pathologies.

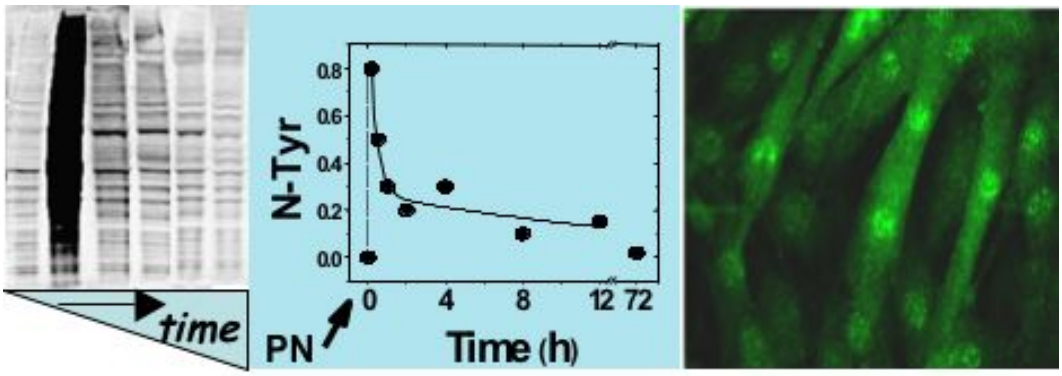


Figure 1. Loss of nitrated proteins in differentiated C2C12 myotubes after exposure to 250-μM peroxynitrite (PN). Myocytes were lysed at various times after treatment; proteins were separated by SDS-polyaerylamide gel electrophoresis (SDS-PAGE) and immunoblotted with anti-nitrotyrosine antibody (left); lanes from left to right represent increasing times after PN addition: 0, 10 min, 12 h, 24 h, 36 h, 48 h, and 72 h. Center panel: The density associated with the extent of nitrotyrosine immunoreactivity plotted as a function of time after PN addition (arrow). The right panel shows PN-treated myotubes (10 min after treatment) fixed and immunostained with anti-nitrotyrosine antibody and secondary antibody conjugated with Alexa-488 dye (green). Nitration is diffuse in the cytoplasm with nucleolar concentration.

Appropriate cellular response to either acute or chronic oxidative stress is likely to be mediated by the modification of sensitive proteins that may act as sensors. Current methods for detecting nitrated proteins are limited to high-abundance proteins. Therefore, we have used the ultra-high-sensitivity mass spectrometry (MS) facility at the W.R. Wiley Environmental Sciences Laboratory to identify proteins most sensitive to nitration. Myocytes were exposed to the precursor molecule SIN-1, which continuously generates peroxynitrite under conditions that result in the slow accumulation of levels of nitrated proteins slightly over baseline levels. The resulting treated and control cells were lysed and treated identically with reduced and alkylated cysteine residues. These samples were subjected to tryptic digestion with subsequent removal of detergent and salts with strong cation exchange chromatography prior to microliter liquid chromatography tandem mass spectrometric (LC-MS/MS) analyses. Ten LC-MS/MS runs were performed per sample, each focusing on a portion of the total 400 to 2000 mass charge (m/z) range covered. The SEQUEST algorithm was configured to search for both native and nitrotyrosine-containing peptides.

A total of 1742 unique proteins were identified, and of these, 94 (or 5%) were nitrated. To mine this wealth of information, the data were organized using the Gene Ontology program that groups identified proteins by both sub-cellular location and associated biological processes. This analysis shows protein recovery from most cellular compartments; nitrated proteins show no preferential distribution with respect to sub-cellular location. However, with respect to biological processes, the nitrated proteins show more than a twofold preference for the category 'response to stimulus,' which includes subcategories of 1) response to stress, 2) response to endogenous stimuli, and 3) response to external stimuli. Among these, the top five nitration targets are all associated with inflammatory and antioxidant responses.

The unanticipated presence of nitration-sensitive proteins within the inflammatory pathway suggests that nitration may provide a way to inhibit this signaling pathway, which

up-regulates the production of nitric oxide, a precursor of peroxynitrite. Moreover, previous studies show that an inflammatory response is upregulated in aging muscle, suggesting that nitration may provide an additional dimension of regulation that has not been considered to date. This work represents development of new approaches testing the applicability of global proteomic screens for identification of post-translationally modified proteins of low abundance, which are likely to be important for elucidating biological stress responses.

This work is an important step on the road to a systems-biology understanding of how biological systems function and interact. Whole-protein catalogs that can identify proteins modified in low copy number give us the ability to correlate responses among groups or families of proteins that will allow us to fill in the blanks in the biochemical networks that are activated and respond to stimuli. Most of the processes that take place in a cell are driven by complexes of several proteins, peptides, and nucleic acids and only more global approaches will allow for a sufficiently comprehensive picture to understand how these systems function.

The work is funded by the National Institutes of Health's National Institute on Aging and Pacific Northwest National Laboratory's Biomolecular Systems Initiative. These results were presented at two recent invited conferences: the 4th International Conference on Peroxynitrite and Reactive Nitrogen Species in Biology and Medicine, July 27-31, 2004, University of Konstanz, Germany; and the Oxidative Post-Translational Modifications in the Cardiovascular System (OPTM) Symposium, October 6-8, 2004, Boston University School of Medicine.

Comparative Proteomics and Cytochrome c Location Determination in *Geobacter sulfurreducens*

KK Hixson,[(a)] YR Ding [(b)] MS lipton,[(a)] and D Lovley[(b)]

(a) Pacific Northwest National Laboratory, Richland, Washington

(b) University of Massachusetts, Amherst, Massachusetts

The study of the detection and location determination of c-type cytochromes in Geobacter sulfurreducens *may lead to an understanding of mechanisms related to electron transport fundamental to bioremediation and energy production.*

Geobacter sulfurreducens is an anaerobic metal-reducing soil bacteria that possesses beneficial capabilities ranging from bioremediation through the reduction of heavy metals and radionuclides to electricity production through electron transport to electrodes on its outer surface(Bond and Lovley 2003; Lovley 1995). The genome sequence of *G. sulfurreducens* reveals a number of novel capacities that have not been previously characterized (Methe et al. 2003). To understand the mechanisms involved in bioremediation and electricity production it is first necessary to determine the proteins that are expressed under different physiological conditions in different locations within the cell. To this end the global proteome of *G. sulfurreducens* was characterized by analysis of proteins from cells grown under eight different conditions with ultra-high-pressure liquid chromatography, and mass spectrometry using the accurate mass and time tag methodology. Cluster analysis using OmniViz™ revealed the reproducibility of triplicate analyses and relevant protein abundance changes among cells grown with different electron acceptors and environmental stress. 2968 gene products (approximately 86% of total putative open reading frames in the genome) were detected including 91 putative c-type cytochrome gene products. Comparative cluster analysis using normalized abundance values showed relevant proteins expressed under eight different environment culture conditions including perturbations of heat shock and oxygen exposure (Figure 1). Highly abundant proteins expressed under iron (III) reduction included cytochromes c (OmcB, PpcE, OmcS and MacA), a glycosyltransferase, lipoproteins, a metal efflux protein, an ABC transporter, an extracellular solute-binding protein, signal transduction and response regulators, and conserved hypothetical proteins. Proteins specifically related to oxygen reduction for energy conservation and heat-shock response were additionally identified.

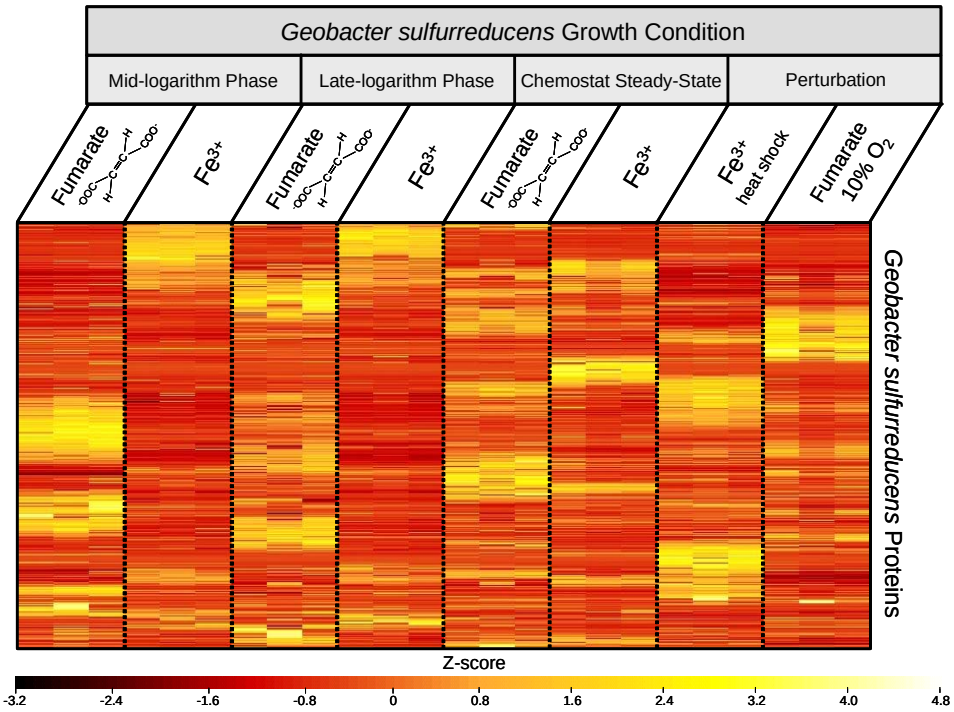


Figure 1. Cluster analysis of 2968 proteins expressed in eight different culture conditions. Lighter-colored areas represent proteins expressed in a higher-than-average abundance because of that culture condition.

Fundamental to this organism's metabolism and metal-reducing capabilities are the transport of electrons from the inner cytosol to the external bacterial surface. At the heart of electron transport is a chain of proteins with heme moieties, which allow for the movement of electrons from one protein to another. These proteins are classified as cytochromes. Unique in this organism is the genetic coding of more than 110 putative c-type cytochrome genes (Methe et al. 2003). The identity and location of cytochrome c gene products were determined by first separating the inner and outer membranes by sucrose gradient centrifugation and subsequently identifying the proteins and performing a cluster analysis, which revealed the predominate location of 91 c-type cytochromes (Figure 2). This information may reveal the paths of fundamental electron transport between proteins in *G. sulfurreducens*, thus revealing the mechanisms involved in bioremediation and electricity production.

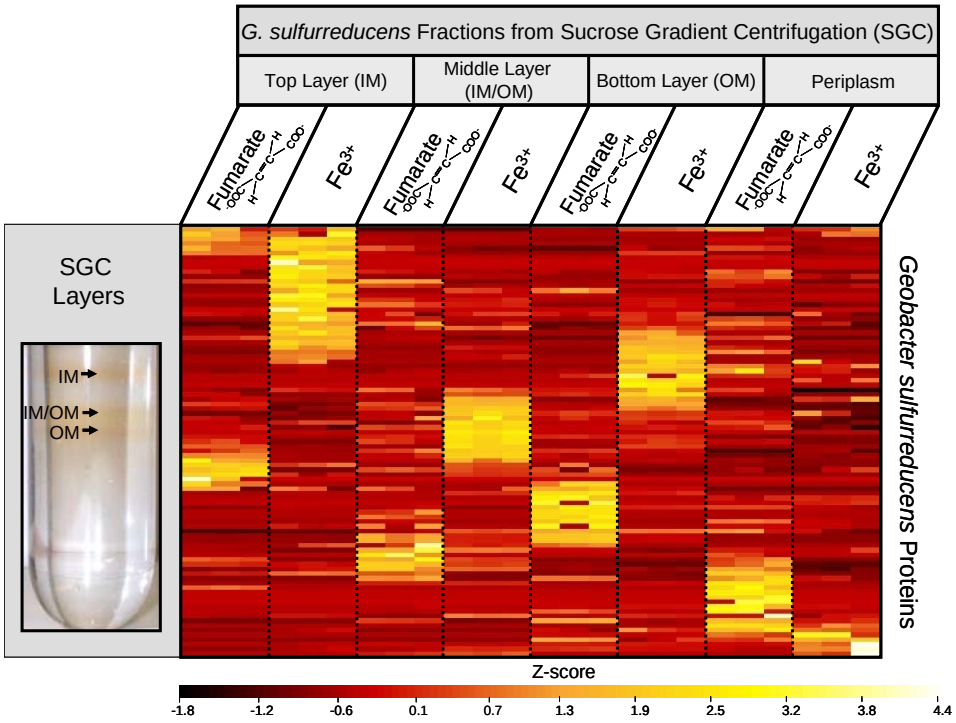


Figure 2. Cluster analysis of 91 c-type cytochromes revealing the predominate location of these proteins from proteins isolated into inner, outer, and periplasmic locations within the bacteria *G. sulfurreducens*. Lighter-colored regions revealed proteins that show a greater-than-average normalized abundance level for that culture condition.

References

Bond DR and DR Lovley. 2003. "Electricity Production by *Geobacter Sulfurreducens* Attached to Electrodes." *Applied and Environmental Microbiology* 69(3):1548-1555.

Lovley DR. 1995. "Bioremediation of Organic and Metal Contaminants with Dissimilatory Metal Reduction." *Journal of Industrial Microbiology* 14(2):85-93.

Methe BA, KE Nelson, JA Eisen, IT Paulsen, W Nelson, JF Heidelberg, D Wu, M Wu, N Ward, MJ Beanan, RJ Dodson, R Madupu, LM Brinkac, SC Daugherty, RT DeBoy, AS Durkin, M Gwinn, JF Kolonay, SA Sullivan, DH Haft, J Selengut, TM Davidsen, N Zafar, O White, B Tran, C Romero, HA Forberger, J Weidman, H Khouri, TV Feldblyum, TR Utterback, SE Van Aken, DR Lovley, and CM Fraser. 2003. "Genome of *Geobacter Sulfurreducens*: Metal Reduction in Subsurface Environments." *Science* 302(5554):1967-1969.

Comparative Proteomics of *Yersinia pestis*

KK Hixson,[(a)] JN Adkins,[(a)] ME Monroe,[(a)] SL McCutchen-Maloney,[(b)] and MS Lipton[(a)]
(a) Pacific Northwest National Laboratory, Richland, Washington
(b) Lawrence Livermore National Laboratory, Livermore, California

The comparative proteome changes detected in Yersinia pestis *lead to quick identification of potentially useful biomarkers to aid in detection, post-infection treatment, and vaccine development for this lethal pathogen.*

The ability to respond adequately to a biological threat requires detailed understanding of the threat agent to determine the appropriate response. Proteins serve as the basis of all mechanisms in biological systems, and understanding the changes in protein expression pattern is crucial to understanding their physiological mechanisms. *Yersinia pestis*, the causative agent of plague, is a known bioterrorist and biowarfare threat agent (Perry et al. 1997). Virulence-associated proteins of *Y. pestis* can be induced or repressed in culture by adjusting conditions that mimic the environments of the flea vector (26ºC) and the human host (37ºC with and without calcium). Insight into the virulence mechanisms of *Y. pestis* can therefore be furthered by the characterization of protein expressed under contrasting temperatures and calcium concentration.

Our approach to global proteomics is based on protein digestion followed by peptide identification by mass spectrometry using the accurate mass and time (AMT) tag methodology (Smith et al. 2002). This approach provided comparative normalized abundance values for 962 proteins identified with two or more unique peptides (Figure 1).

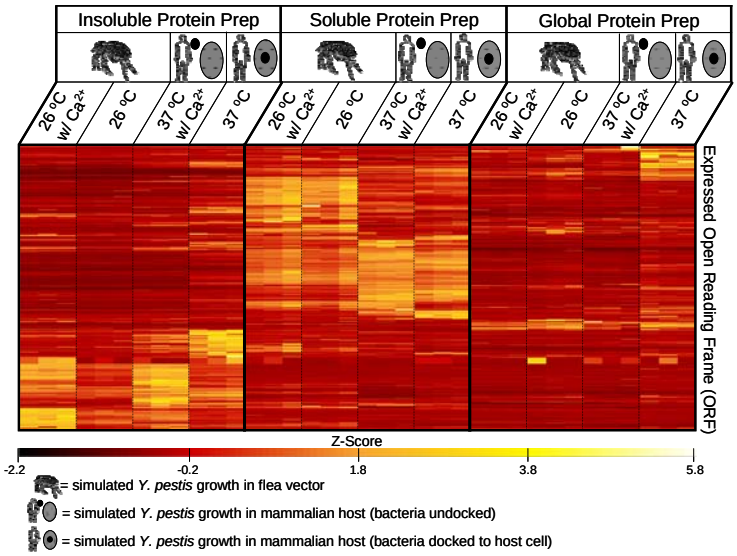


Figure 1. Comparative cluster analysis of expressed proteins (open reading frames, or ORFs) detected using abundance measurements from liquid chromotography-mass spectrometry-Flourien transform ion cyclotron resonance (LC-MS-FTICR) analysis. Lighter colors highlight proteins that are higher in abundance than the average for that culture condition. Darker colors highlight proteins that display a lower-than-average abundance for that culture condition.

Furthermore, cluster analysis was performed on 146 proteins identified from similarity clusters of Figure 1 that displayed similar abundance patterns as nine known virulent proteins (Figure 2). These results have revealed reproducible measurements between triplicate analyses with known and novel protein abundance changes by conditions that mimic virulent and non-virulent bacterial states. Virulent protein expression is clearly observed between bacterial growth from 26ºC to 37ºC with calcium (bacteria's initial entry into mammalian host) and at 37ºC without calcium (bacterial docking with host cell).

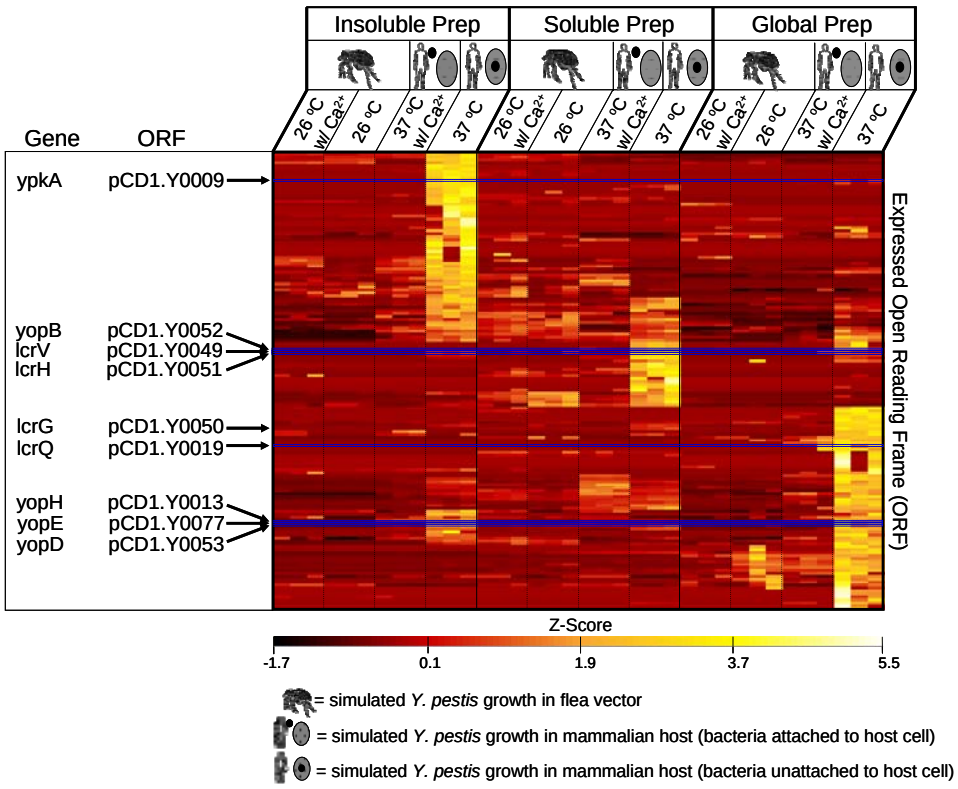


Figure 2. Cluster analysis of 146 proteins that show similar abundance changes to 9 known virulent proteins.

Many of the proteins that were found to cluster in a similar pattern with the nine known virulent proteins shown in Figure 2 were novel hypothetical proteins from both the chromosome and plasmids. It is these novel proteins that most likely also play a role in virulence. These newly identified proteins thus make perfect candidates as biomarkers for protein-based detection technologies and better post-infection treatment and vaccine development.

References

Perry RD and JD Fetherston. 1997. "*Yersinia pestis*-etiologic agent of plague." *Clinical Microbiology Reviews* 10(1):35-66.

Smith RD, GA Anderson, MS Lipton, L Pasa-Tolic, Y Shen, TP Conrads, TD Veenstra, and HR Udseth. 2002. "An Accurate Mass Tag Strategy for Quantitative and High-Throughput Proteome Measurements." *Proteomics* 2(5):513-523.

Global Analysis of the Response of Human Mammary Epithelial Cells to Epidermal Growth Factor Stimulation

T Liu,[a] KM Waters,[(a)] LE Kathmann,[(a)] DG Camp,[(a)] RD Smith,[(a)] HS Wiley,[(a)] and BD Thrall[(a)]

(a) Pacific Northwest National Laboratory, Richland, Washington

Understanding the response of human mammary epithelial cells to epidermal growth factor stimulation can lead to early detection of breast cancer and prediction of its spread, which are critical to successful treatment. This is the first multi-parameter analysis of the response of human mammary epithelial cells to epidermal growth factor stimulation. These data will be used to generate cell-response models, which will aid in our understanding of how multiple signaling pathways interact in a coordinated manner to dictate the cellular response.

One of the most challenging goals of systems biology is to understand the underlying network structure of mammalian cells and how a specific input is translated into an appropriate output. Some progress has been made in building computational models of signaling networks, but the high-resolution and comprehensive datasets necessary to test such models do not exist. To address this issue, we are developing a comprehensive inventory of the genes and proteins expressed in synchronized human mammary epithelial cells (HMEC) during the G1-S transition initiated by epidermal growth factor (EGF) addition. Our current database contains more than one million data points from eight different time points over a 24-hour period, and includes whole genome gene expression profiles. It also includes identification of nearly 7000 intracellular and extracellular proteins by high-throughput mass spectrometry (Figure 1) (Liu et al. 2004; Smith et al. 2003) and western blot analysis using approximately 1000 antibodies. We found that the addition of EGF causes dynamic changes in the expression of more than 700 genes. In addition to the expected changes in cell cycle regulatory proteins, we observed a dynamic set of responses, such as down-regulation of growth inhibitory genes, enhanced expression of epidermal growth factor receptor (EGFR) ligands, and induction of many cytokine signaling pathway components. In addition, the mitogenic response to EGF was associated with a marked increase in proteins that build and modify the extracellular matrix, including matrix metalloproteases. This suggests that the cellular decision to proliferate is tightly coupled to remodeling the local microenvironment. These data are being used to create a comprehensive network map of HMEC and realistic models of signal transduction pathways.

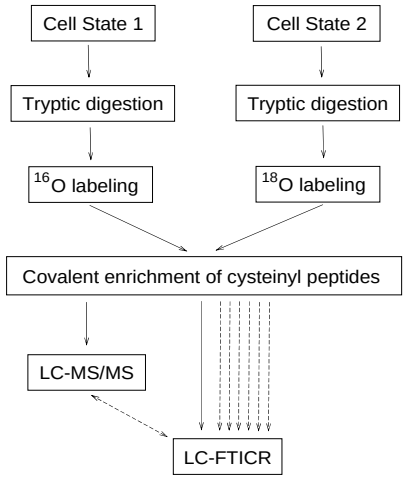


Figure 1. Strategy for quantification of differential protein expression using Quantitative Cysteinyl-peptide Enrichment Technology (QCET).

The EGFR plays an important role in regulating the growth and behavior of normal epithelial cells (Stampfer et al. 1993; Dong et al. 1999). Previous studies in normal HMECs have shown that both proliferative growth and cell motility are highly dependent on EGFR signaling. In addition, a growing number of studies indicate cellular responses to G-protein-coupled receptor ligands, cytokines, and many chemicals are mediated in part through secondary activation (transactivation) of EGFR, through direct coupling of intracellular signaling molecules, induced shedding of EGFR ligands, or interactions with matrix complexes such as integrins (Carpenter 1999; Chen et al. 2004; Fischer et al. 2003). Such reports suggest an important function of the EGFR pathway is to integrate cellular signaling responses from a wide range of receptor-mediated pathways, thus modulating signaling in a context-dependent manner. To understand this integrating function of the EGFR pathway however, it is necessary to define the pathways that are directly initiated by EGFR activation. Thus, the goal of this project is to develop a comprehensive inventory of signaling pathways stimulated by EGFR. Because of the important role that EGFR plays in regulating mitogenic responses of HMEC, we have focused our initial studies on defining the mRNAs, proteins, and kinase pathways activated during the mitogenic response (G1-S phase transition) to EGFR activation. Ultimately, these data will be used to generate cell-response models that will aid in our understanding of how multiple signaling pathways interact in a coordinated manner to dictate the cellular response.

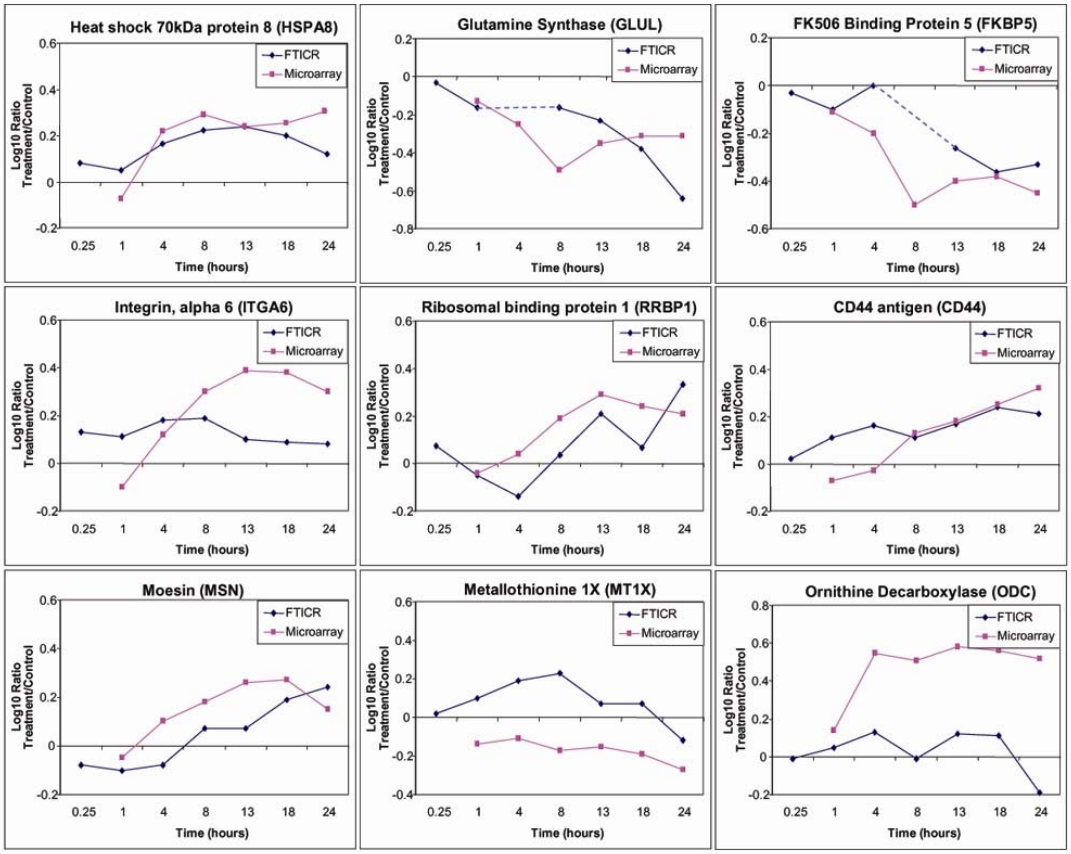


Figure 2. Comparison of expression ratios obtained by microarray and Fourier transform ion cyclotron-mass spectrometry (FTICR-MS) analysis.

The FTICR-MS approach provided approximately tenfold greater identification of protein changes as did the Powerblot approach. However, each approach has advantages. Phosphorylation data provided by the Powerblot analysis improved the confidence in the pathway analysis, whereas the FTICR-MS is more robust in identification of large numbers of protein changes. A comparison of the proteomic data and the genomic data has also been made (Figure 2).

The integrated results obtained in this study suggest that the major early pathways stimulated by EGFR involve Src, ERK, p38, STAT3, and NF-kB as signaling mediators (Figure 3). These results are consistent with the known literature. Genes thought to be required for the G1-S transition were also found to be induced in a temporally consistent manner, providing confidence that the expression results obtained are representative of the cell response. Among the most highly induced responses arising at later time points included the induction of several metalloprotease pathways associated with modification of the extracellular matrix. Whether the mitogenic response and the matrix modifying response of EGFR signaling are tightly coupled, or behave as independently regulated pathways requires further study.

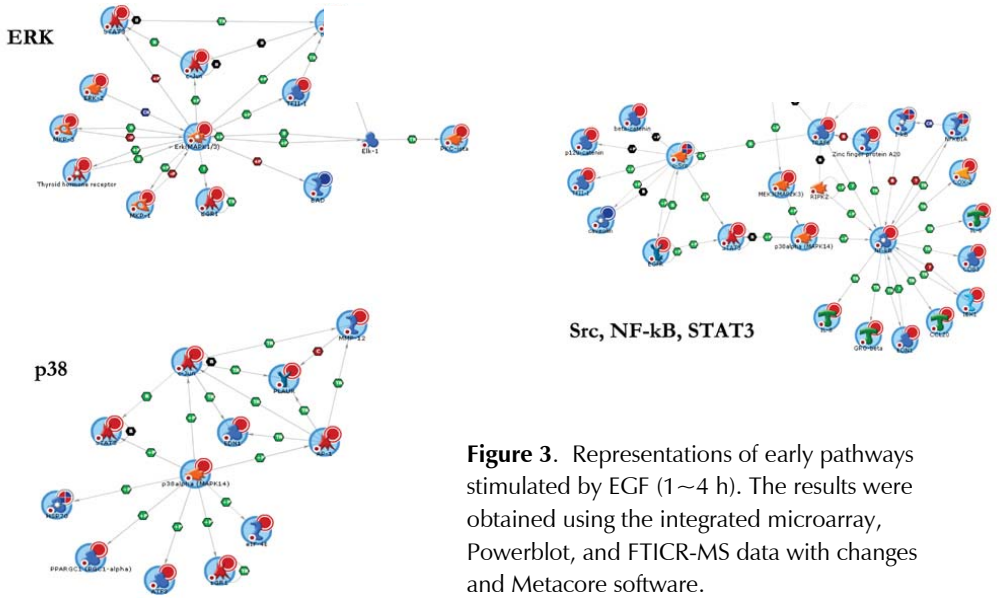


Figure 3. Representations of early pathways stimulated by EGF (1~4 h). The results were obtained using the integrated microarray, Powerblot, and FTICR-MS data with changes and Metacore software.

Only a fraction (25 to 30%) of cells undergo a mitogenic response under the conditions used in this study. Nonetheless, we were able to discern gene expression changes that correspond to the G1-S transition in the whole cell population. Such results suggest that the responses we have measured may be dampened in the overall population compared to the 'responding' population. These results highlight the need for imaging-based approaches, which permit quantitation of signaling responses in individual cells to investigate whether the pathways identified are activated in all cells, or only in subsets of responding cells.

References

Carpenter G. 1999. "Employment of the Epidermal Growth Factor Receptor in Growth Factor–Independent Signaling Pathways." *Journal of Cell Biology* 146(4):697-702.

Chen W-NU, RL Woodbury, LE Kathmann, LK Opresko, RC Zangar, HS Wiley, and BD Thrall. 2004. "Induced Autocrine Signaling through the Epidermal Growth Factor Receptor Contributes to the Response of Mammary Epithelial Cells to Tumor Necrosis Factor Alpha." *Journal of Biological Chemistry* 279(18):18488.

Dong J, LK Opresko, PJ Dempsey, DA Lauffenburger, RJ Coffey, and HS Wiley. 1999. "Metalloprotease-Mediated Ligand Release Regulates Autocrine Signaling Through the Epidermal Growth Factor Receptor." *Proceedings National Academy Sciences of the United States of America* 96(11):6235-6240.

Fischer OM, S Hart, A Gschwind, and A Ullrich. 2003. "EGFR Signal Transactivation in Cancer Cells." *Biochemical Society Transactions* 31(6):1203-1208.

Liu T, W-J Qian, EF Strittmatter, DG Camp II, GA Anderson, BD Thrall, and RD Smith. 2004. "High-Throughput Comparative Proteome Analysis Using a Quantitative Cysteinyl-peptide Enrichment Technology." *Analytical Chemistry* 76(18):5345-5353.

Smith RD, GA Anderson, MS Lipton, L Pasa-Tolic, Y Shen, TP Conrads, TD Veenstra, and HR Udseth. 2002. "An Accurate Mass Tag Strategy for Quantitative and High-Throughput Proteome Measurements." *Proteomics* 2(5):513-523.

Stampfer MR, CH Pan, J Hosoda, J Bartholomew, J Mendelsohn, and P Yaswen. 1993. "Blockage of EGF Receptor Signal Transdcution Causes Reversible Arrest of Normal and Immortal Human Mammary Epithelial Cells with Synchronous Reenty into the Cell Cycle." *Experimental Cell Research* 208(1):175-188.

Global Quantitative Proteomics Without Isotope Labeling

R Fang,[(a)] DA Elias,[(a)] ME Monroe,[(a)] Y Shen,[(a)] CD Goddard,[(a)] SJ Callister,[(a)] RJ Moore,[(b)] YA Gorby,[(a)] JK Fredrickson,[(a)] MS Lipton,[(a)] and RD Smith[(a)]

(a) Pacific Northwest National Laboratory, Richland, Washington

(b) W.R. Wiley Environmental Molecular Sciences Laboratory, Richland, Washington

Proteomics, the study of the protein product of the genome, increases our understanding of biological pathways and leads to the discovery of biomarkers. Faster and less expensive analyses are needed before routine diagnostic use can be made of proteomics. This work is an important step in that direction as it eliminates expensive isotopic labeling and measures changes in abundance directly from the signals in the mass spectrometer.

We are investigating an alternative approach for the relative quantification of protein abundances utilizing liquid chromatography mass spectrometry (LC-MS) technologies. Current approaches based on stable isotopic labeling of proteins, or components of proteins known as peptides, either chemically or metabolically can be cost-prohibitive and often result in a fewer number of proteins being detected by LC-MS. This alternative approach does not require isotopic labeling, but requires a linear correlation between the amount of the analytes and their peak areas measured by MS. This has been demonstrated with simple mixtures of a few analytes (Tang et al. 2004; Cech and Enke 2001; Voyksner and Lee 1999), and it has been suggested that this linear correlation can be obtained for more complex mixtures of analytes using a sufficiently low flow rate and a small amount of sample. With such conditions, the ionization of analytes using electrospray ionization approaches an optimum (~100% ionization efficiency) (Tang et al. 2004; Cech and Enke 2001).

This alternative approach, abbreviated as AMT-NIQ (accurate mass and time tag non-isotopic labeling quantification), utilizes the previously developed AMT tag procedure for high-throughput confident identification of peptides (Smith et al. 2002; Lipton et al. 2002). We applied the AMT-NIQ approach to the metal-reducing bacteria *Shewanella oneidensis* strain MR-1 cultured in bioreactors under both aerobic (20% dissolved oxygen tension) and suboxic (< 0.1% oxygen) conditions. Arbitrary abundances of peptides from proteolytic digests of proteins extracted from collected samples were measured using liquid chromatography-Fourier transform ion cyclotron resonance (LC-FTICR) operated at a conventional flow rate. Normalization of the peptide abundance across all LC-FTICR analyses to correct for possible systematic bias was performed.

Experimental and instrument reproducibility of the AMT-NIQ approach was demonstrated using scatter plots. An example displayed in Figure 1 (top) plots an instrument replicate of abundances of peptides from one collected sample versus an instrument replicate of abundances of peptides from a duplicate sample. These samples were collected simultaneously while the bioreactor operated under an aerobic condition. We observed good reproducibility among all instrument replicates from these duplicate samples as indicated by root sum square (R^2) values ranging from 0.95 to 0.96. Figure 1 (bottom) plots an instrument replicate from one bioreactor operating under an aerobic condition against an instrument replicate

from a different bioreactor operating under a suboxic condition. Here, R^2 values among instrument replicates were significantly lower, ranging from 0.65 to 0.67, indicating that the abundance of certain peptides – and therefore the proteins they represent – were different between culture conditions.

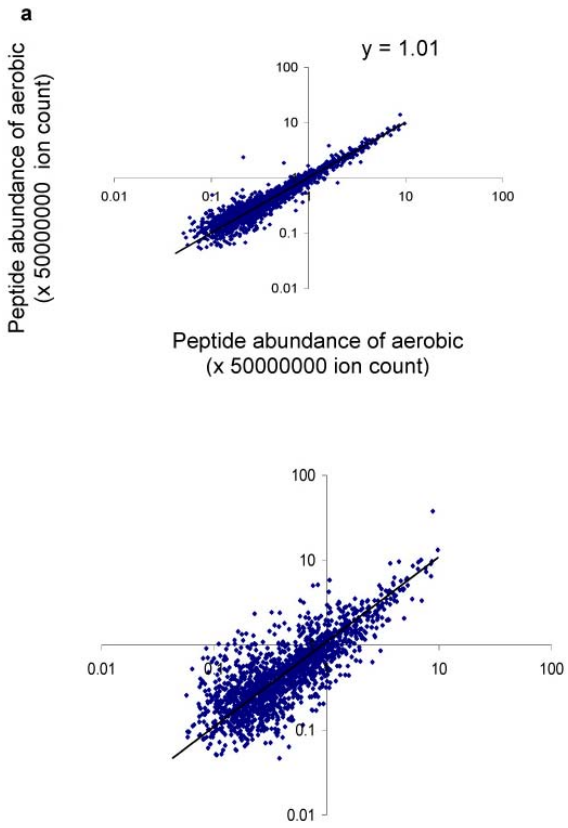


Figure 1. (top) Scatter plot of abundances of peptides from an instrument replicate from one collected sample versus abundances of peptides from an instrument replicate taken from a duplicate sample. Each sample was collected from the same bioreactor operating under an aerobic environment. (bottom) Scatter plot of abundances of peptides from an instrument replicate from one bioreactor operating under an aerobic environment versus abundances of peptides from an instrument replicate from a different bioreactor operating under a suboxic environment.

References

Cech NB and CG Enke. 2001. "Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals." *Mass Spectrometry Reviews* 20(6):362-387.

Lipton MS, L Pasa-Tolic, GA Anderson, DJ Anderson, DL Auberry, JR Battista, MJ Daly, J Fredrickson, KK Hixson, H Kostandarithes, C Masselon, LM Markillie, RJ Moore, MF Romine, Y Shen, E Stritmatter, N Tolic, HR Udseth, A Venkateswaran, KK Wong, R Zhao, and RD Smith. 2002. "Global Analysis of the *Deinococcus Radiodurans* Proteome by Using Accurate Mass Tags." *Proceedings of the National Academy of Sciences of the United States of America* 99(17):11049-11054.

Smith RD, GA Anderson, MA Lipton, L Pasa-Tolic, Y Shen, TP Conrads, TD Veenstra, and HR Udseth. 2002. "An Accurate Mass Tag Strategy for Quantitative and High-Throughput Proteome Measurements." *Proteomics* 2(5):513-523.

Tang K, JS Page, and RD Smith. 2004. "Charge Competition and the Linear Dynamic Range of Detection in Electrospray Ionization Mass Spectrometry." *Journal of the Society for Mass Spectrometry* 15(10):1416-1423.

Voyksner RD and H Lee. 1999. "Investigating the Use of an Octupole Ion Guide for Ion Storage and High-Pass Mass Filtering to Improve the Quantitative Performance of Electrospray Ion Trap Mass Spectrometry." *Rapid Communications in Mass Spectrometry* 13(14):1427-1437.

Illumination of Previously Unidentified Proteins with Specific Enzyme Activity Using Mass Spectrometry

RC Barry,[(a)] DH Kohl,[(b)] DG Camp,[(a)] and RD Smith[(a)]
(a) Pacific Northwest National Laboratory, Richland Washington
(b) Washington University, St. Louis, Missouri

While genomes have been sequenced, the function of most genes and their corresponding proteins are still unknown. It's easier to unravel the complexity of the rules that govern this functionality by starting with simpler organisms, such as yeast. This work is an important step in increasing our ability to identify the unknown functionality of genes when more than one protein or peptide participates in the chemical/enzymatic activity, which may lead to the ultimate correlation between specific activities and participating genes.

Researchers are collaborating to establish a mass spectrometry-based approach for identifying those proteins responsible for specific enzymatic activity, but whose identity was previously unknown. While the availability of complete genome sequences opens the door to important biological advances, much of the actual understanding of cellular systems and the roles of its constituents will be based on techniques in proteomics (which we define here as the study of the functions of proteins). Current gene annotation techniques fall short of accurately identifying the functions of proteins within an organism. Many proteins, as identified by open reading frames, have unknown function and are listed as 'hypothetical' proteins. Further, there are many biological processes that are carried out by organisms whose protein players are completely unknown. Often, many of the enzymes that participate in biochemical pathways are completely unknown and in some cases have been identified from previously classified hypothetical proteins. Developing a rapid and sensitive method for associating enzymatic activities with proteins from a specific gene would greatly accelerate our ability to gain an understanding of the structure and function relationships of enzymes and the biological pathways used by organisms to promote survival.

Present methods typically use a series of purification steps to isolate fractions containing specific enzymatic activity. The proteins from these fractions are then separated with Coomassie Blue-stained bands using sodium dodecyl sulfate polyacrylamide gel electrophoresis. Identification of the enzymes responsible for the activity is inferred from the intensity of the stained bands. Theoretically, the more purification steps one uses, the greater the purity of the compound that is producing the activity and the higher the specific activity detected. Whole-cell lysates, for example, may be fractionated using anion exchange chromatography. Those fractions containing the activity can be isolated and further resolved using size-exclusion chromatography, where again those fractions containing specific activity can be isolated for further analysis. We are studying the reduction of methylglyoxal in yeast cells. Extensive work at Washington University using the conventional approach described above identified the open reading frame that expresses the nicotinamide adenine dinucleotide phosphate hydrogen (NADPH)-dependent enzyme responsible for this activity.

These results are shown in Figure 1 where the open circles show the specific activity for methylglyoxal reductase (NADPH-dependent). This was increased one hundred twenty-twofold going from anion exchange chromatography (upper figure) to size-exclusion chromatography (lower figure). In this preliminary work, the proteins are extracted from each of the fractions that demonstrate increased enzymatic activity, digested with trypsin, and then analyzed using ultra-sensitive liquid chromatography-Fourier transform ion cyclotron resonance (LC-FTICR) mass spectrometry. The concentration-dependence of the resulting peptides, ionized by the electrospray ionization process, are used to determine the identity of the protein(s), which similarly demonstrate an increase in their concentration in solution. Figure 1 shows the results for the enzyme methylglyoxal reductase as the solid circles. They show general agreement with the enzymatic activity and point out more general issues associated with linking protein concentrations with enzymatic activity. Often co-factors are required to activate an enzyme and only a complete inventory of all biologically important molecules present in a fraction can lead to a complete understanding.

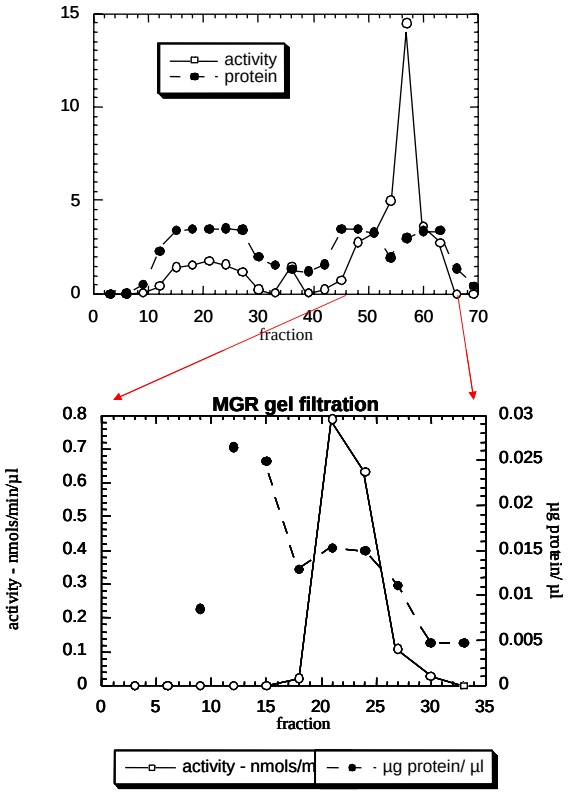


Figure 1. Isolation of methyglyoxyl reductase activity using anion exchange and size-exclusion chromatography.

With the LC-FTICR technology, all other proteins present can also be detected as possible candidates. Once a candidate list is produced, further experiments can be performed to validate protein function, such as using glutahione tranferase (GST)-fusion tags to over express the protein of interest in yeast to provide protein for biochemical assays. Additional experiments are underway to minimize the number of different purification steps required by collecting and analyzing the protein content of the individual fractions from a single purification run to identify the protein(s), which dramatically increase in concentration as well as specific activity.

Initial Progress in the Three-Dimensional Proteomic Mapping of the Mouse Brain

W Qian,[(a)] H Wang,[(a)] VA Petyuk,[(a)] MH Chin,[(b)] DJ Smith,[(b)] DG Camp,[(a)] and RD Smith[(a)]

(a) Pacific Northwest National Laboratory, Richland, Washington

(b) University of California at Los Angeles David Geffin School of Medicine, Los Angeles, California

High-throughput, high-sensitivity proteomics enables a new opportunity for three-dimensional brain proteome mapping. In this study, a mouse brain is sectioned into 600 cubes (voxelations), and the proteins in each cube are detected. Such three-dimensional mapping may, when combined with other maps of brain function and structure, provide novel insights into brain functions and diseases, leading to improved diagnosis and treatment of brain maladies.

Understanding how the genome gives rise to the staggering three-dimensional complexity of the mammalian brain is one of the major goals for post-genomic biology. To address this problem, it will be necessary to spatially map transcripts, proteins, and the networks they form in the brain at a genome-wide level. In this study, we investigated the potential for using high-throughput, high-sensitivity proteomics and the voxelation technology for three-dimensional proteome mapping of the mouse brain. Added value for the use of the mouse model system for three-dimensional mapping arises from the existence of mouse models for a number of neurodegenerative diseases. Figure 1 shows the overall voxelation process. Mouse brain tissue is directly sampled as cubes using a three-dimensional coordinate system. The voxelation was performed by first cutting the brain into 10 coronal sections and then each section was cut into ~60 cubes with each cube of ~0.75 μL (Singh et al. 2003). The voxels were directly put into 96 well plates for automated sample processing and high-throughput proteome analyses.

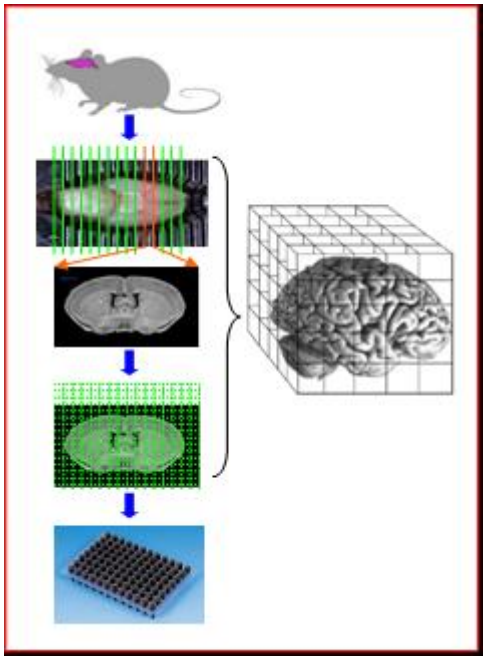


Figure 1. Schematic of the voxelation process. The mouse is sacrificed and the brain extracted. The brain is sectioned into layers and each layer is then cubed by a rectangular cutting grid.

The final phase of peptide/protein identification will use high-throughput liquid chromatography-Fourier transform ion cyclotron resonance (LC-FTICR) analyses based on the accurate mass and time (AMT) tag approach. This approach requires the generation of AMT tag database as a first step and as validation of the approach. We have completed the first phase validating the approach and have generated an extensive AMT tag database for

the mouse brain. This required the development of novel and highly efficient enrichment techniques for cysteine-containing peptides. The samples were then fractionated and analyzed using multidimensional LC followed by online tandem mass spectrometry (MS/MS) detection using ion trap mass spectrometers. The database currently contains more than 35,000 different AMT tags, covering ~6900 different mouse brain proteins and is the most extensive protein database ever generated for the mouse brain. The proteins detected were sorted according to the subcellular component in which they would be found according to annotations in the genome. Figure 2 shows the coverage of identified brain proteins based on subcellular components. The complete coverage of mouse brain subcellular locations in the AMT tag database provides a solid basis for future quantitative analyses of brain voxels.

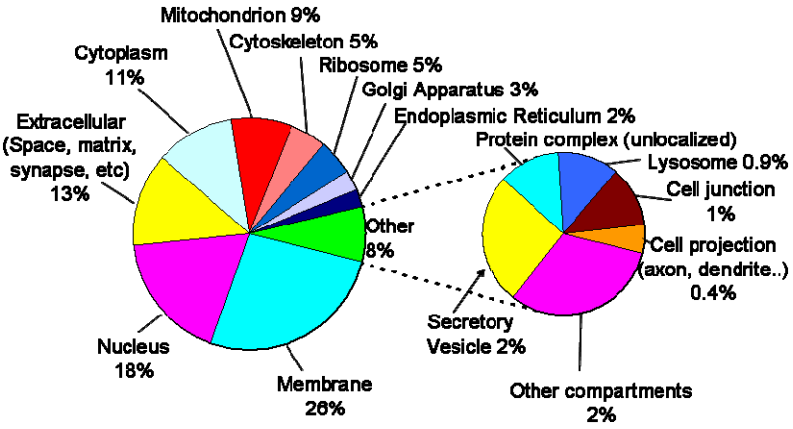


Figure 2. Distribution of proteins in our database amongst the subcellular components according to annotations of the mouse genome.

Success of the next phase of the project requires the analyses of single voxel samples using an automated micro-scale sample process and extremely sensitive LC-MS detection using W.R. Wiley Environmental Molecular Sciences Laboratory's FTICR mass spectrometers. These requirements have presented procedural challenges that have been overcome. We have developed a simple protocol for tissue lysis and protein digestion, and initial automated sample processing was successful. With this protocol, the typical amount of peptides recovered per voxel was ~20 μg. With our in-house developed capillary LC separations, this is enough for both LC-MS/MS and LC-MS analyses. Figure 3 shows the successful analyses of two adjacent single voxels using LC-FTICR. For each voxel approximately 3000 peptides and 1000 proteins were identified using AMT tags from our database. The automated sample processing is still being optimized and improvements in the number of proteins detected per voxel are expected. The next step will be quantitative analyses of large number of voxels for three-dimensional mapping. The normalization and quantitation approaches are still being refined to optimize the confidence of the identification and minimize the error bars on the protein quantities. We anticipate that the initial generation of three-dimensional protein abundance patterns (e.g., for normal and Parkinson disease mouse models) will provide novel insights.

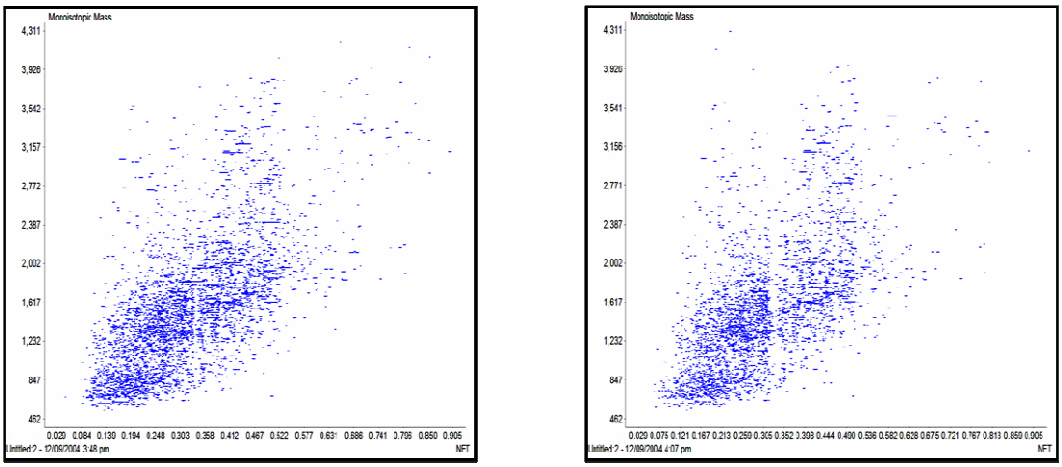


Figure 3. Two-dimensional-plots for the LC-FTICR analysis from two neighboring single voxels.

Reference

Singh RP, VM Brown, A Chaudhari, AH Khan, A Ossadtchi, DM Sforza, AK Meadors, SR Cherry, RM Leahy, and DJ Smith. 2003. "High-Resolution Voxelation Mapping of Human and Rodent Brain Gene Expression." *Journal of Neuroscience Methods* 125(1-2):93-101.

Probing the Function of Proteins by Multidimensional Mass Spectrometric Analysis of Breast Cancer Membrane Proteomes

AJ Patwardhan,[(a)] EF Strittmatter,[(b)] DG Camp,[(b)] RD Smith,[(b)] and MG Pallavicin[(a)]

(a) University of California San Francisco, San Francisco, California

(b) Pacific Northwest National Laboratory, Richland, Washington

Early detection and proper identification of the state of a cancer is vital to its effective treatment. Breast cancer is especially critical when determining the state of the disease to predict its progression as an aid to treatment. Here we have identified a large collection of proteins associated with the cell membrane of breast cancer cell lines. Several of them have been identified as possibly playing important roles in the initiation and progression of breast cancer.

Proteins associated with plasma cell membranes participate in multiple cell-environment interactions including cell-cell communications, cell-extracellular interactions, locomotion and migration, and cell signaling. Membrane-associated proteins account for 70% of all known pharmaceutical drug targets and are targets for a number of new drug and antibody-based therapeutics in cancer. We undertook a study to define the membrane proteome of breast cancers. To capture the maximum number of breast cancer-associated proteins, we analyzed the membrane proteome of a panel of breast cancer cell lines using a combination of subcellular fractionation, separation by strong cation exchange chromatography, and reversed-phase capillary liquid chromatography (RPLC) followed by tandem mass spectrometry (MS/MS). The proteins identified were investigated further in an effort to identify plasma membrane-associated proteins (PMAPs) and groups of proteins over-represented in the breast cell-line panel. Additional information derived from gene expression data sets were used to identify proteins that may have a potential relevance to breast cancer.

Four breast cancer cell lines (SKBR-3, MDA-MB-231, BT-474, and an MCF7-c18 transfectant cell line over-expressing the ERBB2 receptor tyrosine kinase [RTK]) were cultured and used to make up the breast cancer panel. Membrane-enriched fractions were prepared separately for all four cell lines and then pooled into a single mixture. All sample aliquots combined to produce 202 fractions, which produced 724,566 MS/MS spectra. This corresponded to a total of 42,250 high-scoring peptides matching 2344 proteins listed in the September 2003 freeze of the human International Protein Index (IPI) database (http://www.ebi.ac.uk/IPI/IPIhelp.html). Among all proteins identified, 28% had less than 10% sequence coverage, 65% had 20 to 50% sequence coverage, and 6% had sequence coverages greater than 60%. Membrane-bound proteins were further subcategorized into cellular organelles based on the gene ontology (GO), UniProt, and InterPro databases (Figure 1) and ambiguities were resolved using PSORTII.

The transmembrane-spanning region of proteins is frequently lost in membrane-enriched preparations because of the inability to keep hydrophobic regions of the protein solubilized. A decrease in the number of transmembrane domains (TMDs) may suggest that portions of the membrane protein embedded within the lipid bilayer are not captured efficiently with our

isolation procedure or analytical strategies. The number of TMDs was predicted using the software package TMHMM 2.0 prediction server and was used to extrapolate the number of TMDs expected for 2198 (94%) of the identified proteins. Percent protein coverage varied greatly among both membrane containing and membrane-free proteins and no significant relationship between the number of TMDs and percent protein coverage was observed. However, sequence coverage above 40% was generally not seen in proteins with multiple TMDs. These data suggest that there may be some bias introduced against hydrophobic portions of the membrane that may have been lost during sample preparation or analysis. The majority of proteins (66%, 1455) had 0 TMDs, while 34% (743) had between 1 and 23 TMDs as predicted by the TMHMM 2.0 server. Given that global genomic analysis predicts that 20 to 30% of all open reading frames encode integral membrane proteins (including both plasma and intracellular membranes), our results indicate that any bias against membrane proteins is relatively small.

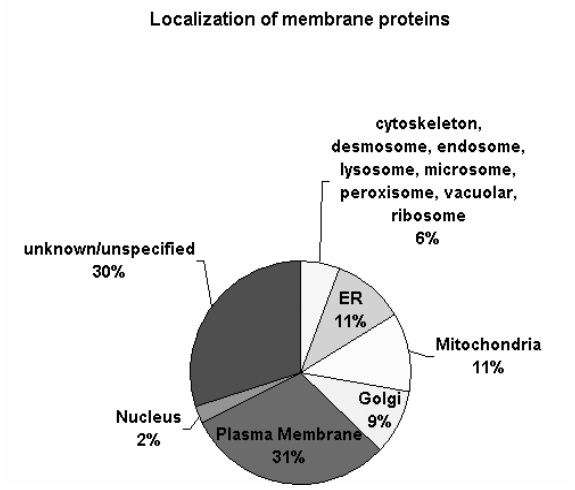


Figure 1. Summary of localization information for 965 membrane proteins identified from MS/MS.

Querying of the UniGene database revealed more than 1905 unique genes represented by this set of proteins. Of these unique genes (identified from all 3 aliquots), 1525 (80%) were mapped to the GO database and classified according to molecular function and cellular component categories. A combination of both molecular function and cellular component categories yielded 312 proteins that were associated with the plasma membrane, as evidenced by both functional and localization information. Because GO categories are often incomplete and ambiguous, several proteins were manually inspected for association with the plasma membrane. Manual inspection revealed an additional 40 PMAPs. Most proteins in the subset of 352 PMAPs were associated with signal transducer activity, binding, and transporter activities. The sub-categorization of our data set according to the GO database resulted in many proteins putatively involved in immune response, apoptosis, catalytic activity, and several other cellular processes. According to the PFAM database, twenty protein sequence domain/families were well represented (more than five member proteins) in our collection of PMAPs. Many of these domains include well-conserved transmembrane- bound domains and included several groups of proteins known to be membrane-associated. Among those identified was the immunoglobulin domain (~100 residues); the fibronectin type III domain (~100 residues); the epidermal growth factor (EGF)-like domain (30-40 residues); the SH3 domains (~50 residues); the WD domain G-beta repeat (~40 residues); the FERM (~150 residues) domain; the PDZ domains (80-90 residues); the tubulin domains; the sterile alpha motif (more than 70 residues); the cadherin domain; and the low-density lipoprotein receptor domain. Well-represented protein families

included members of the ADP-ribosylation factor family, the ABC transporters, the Ras family of green fluorescence protein (GTP)ases, the guanine nucleotide binding G-protein family, and the tetraspanin family.

Using the procedure described by Hosack et al. (2003), protein groupings based on functional domains in several protein signature databases were used to determine whether any populations of proteins were abundant in our sample. Two protein categories (as defined in the PFAM and SMART) contained a significantly ($p<0.01$) higher number of proteins in our sample than would be expected by random chance. Four proteins belonging to the SNF-7 homologues (PFAM, PF03357) domain were identified in our sample. This family of proteins is involved in formation of multivesicular bodies, which play a critical role in the recycling and degradation of membrane proteins (i.e., activated cell-surface receptors). Thirteen proteins belonging to the PDZ/DHR/GLGF (SMART, SM00228) domain were also identified. Members of this group of proteins are thought to direct the targeting of signaling molecules to sub-membranous sites. The PDZ/DHR/GLGF domain group includes proteins frequently associated with the plasma membrane, a compartment where high concentrations of phosphatidylinositol 4,5-bisphosphate (PIP2) are found. They interact directly with several ephrin-related RTKs, the ERBB2 RTK, and with C-terminal sequences derived from other transmembrane receptors.

SAGE tag numbers directly reflect the abundance of mRNAs *in vivo* and can therefore be used as an accurate and quantitative measure of global gene expression profiles in breast tissue. Using archived, publicly available SAGE data sets, we analyzed differences in gene expression between pools of normal and cancerous breast tissue. Differential gene expression is inferred from the relative frequency of tags occurring in pools of cancerous SAGE libraries vs. pools of normal SAGE libraries. The expression profiles for 34 PMAPs showed at least a threefold difference in expression between pools of normal and cancerous breast tissue. Several of the proteins over-expressed in cancerous breast tissue have previously been investigated with regard to breast cancer. The RTK protein ERBB2 has been well documented as a prognostic marker in breast cancer and was over-expressed in all the cancerous tissues surveyed compared to normal tissue. Notably, the myristoylated alanine-rich C-kinase substrate protein, putatively involved in the activation of the ERBB2 signaling pathway, also was over-expressed in tumor tissue. Other over-expressed proteins in tumor tissue included the peripheral benzodiazepine receptor, the ephrin receptor EPHBB3, the melanoma-associated antigen D2, the beta platelet-derived growth factor receptor, the semaphorin 3C receptor precursor, dystroglycan, and Claudin-3. Several of the PMAPs over-expressed in breast cancer tissue had little or no previous work done in relation to breast cancer. These proteins include the tyrosine-protein kinase-like 7 precursor, the EFG receptor kinase substrate, the FK506-binding protein 3, the calcium-binding protein p22, the serine/threonine-protein kinase 25, the Erythrocyte band 7 integral membrane protein, the Cadherin EGF receptor, the retinoic acid-induced 3 protein, the DKFZp761D0211 hypothetical protein, and the GDNF family receptor alpha 1 precursor.

Among the proteins under-expressed in cancerous compared to normal breast tissue are several that show a surprisingly large disparity between both pools. The Annexin A1 protein is under-expressed in all cancerous tissues compared to normal. This protein belongs to a family of calcium-dependent, phospholipid-binding proteins that are preferentially located

on the cytosolic face of the plasma membrane and is thought to have potential anti-inflammatory activity and act as endothelial adhesion molecules. Other proteins under-expressed in cancer tissue include the CD44 and CD63 antigens, the GTP-binding protein, the jagged-1 protein, and the ephrin receptor EPHA2.

In summary, our study identified a large collection of plasma membrane-associated proteins from a breast cancer cell-line panel using high-throughput tandem mass spectroscopy techniques. Several tools allowed us to create a membrane proteome profile of this set of cell lines. In addition, we were able to combine gene expression data with proteomic data to identify several PMAPs that may play important roles in the initiation and progression of breast carcinomas. Because we analyzed a limited number of specimens, additional experiments using high-throughput techniques, such as mRNA *in situ* hybridization or immunohistochemical analysis on tissue-microarrays, are required to determine how commonly these genes are differentially expressed. Additional analysis of these genes and the biochemical pathways in which they are involved will not only further our understanding of breast oncogenesis, but will also provide new and valuable targets for translational research.

Reference

Hosack DA, G Dennis, BT Sherman, HC Lane, and RA Lempicki. 2003. "Identifying Biological Themes Within Lists of Genes with EASE." *Genome Biology* 4(10):R70.

Proteomic Analysis of a Hepatitis C Virus Cell Model System

JM Jacobs[(a)] and DL Diamond[(b)]

(a) Pacific Northwest National Laboratory, Richland, Washington

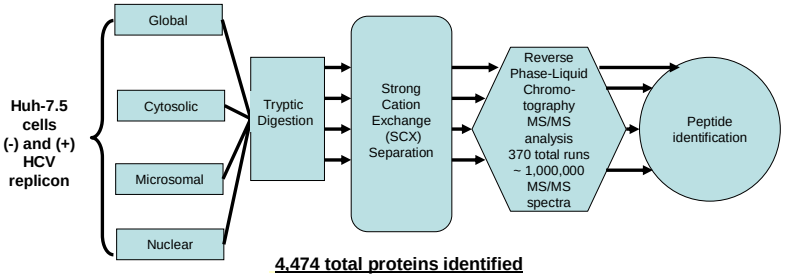
(b) University of Washington, Seattle, Washington

Early and accurate detection of infections is often a key to effectively treating them. A critical first step in developing assays based on the detection of proteins that are characteristic of a specific infection is to determine which proteins are expressed only or predominantly in the infected state. This work is an initial catalogue of proteins that will help characterize the progression of Hepatitis C virus in liver tissue.

Hepatitis C virus (HCV) is the most common blood-borne infection in the United States, infecting approximately 2% of the population. Approximately 85% of cases progress to chronic infection, which often results in liver disease, including variable degrees of hepatic inflammation and fibrosis, cirrhosis, and hepatocellular carcinoma. Since its discovery and sequencing in 1989, studies have been performed to elucidate the host/virus interactions pertinent to persistent infection of the liver. To date, global characterization of the host cellular response to infection has centered on the use of expression microarray profiling to identify potential gene markers of HCV-associated liver disease. In contrast, only limited studies describing proteomic analysis of human liver proteins have been reported. Proteomic studies of HCV infection have been particularly limited for several reasons including the lack of a good cell-culture model and the need for large amounts of protein for conventional proteomic analysis.

Recent advances in both proteomic methodologies and cell-culture models of HCV infection now make it possible to perform global characterization of the host cell protein response within the context of the complete set of HCV genes *in vitro*. Refined multidimensional liquid chromatographic (LC) separations coupled with mass spectrometry (MS) for proteome analysis has allowed global experiments to be performed utilizing less protein, and obtaining more sensitivity, throughput, and dynamic range than with previous proteomic techniques. Concurrently, the establishment of a highly permissive Huh-7 subline (Huh-7.5), which supports high-level replication of full-length HCV genomes, allows for the analysis of host cell response in the presence of persistent viral replication and expression of all viral proteins such that the interactive potential of HCV proteins can be investigated.

Collaborators from Pacific Northwest National Laboratory and the University of Washington have recently performed the first global proteomic analysis of the highly permissive Huh-7.5 cell line in the presence and absence of replication of the full-length HCV genome. Powerful multidimensional separation techniques coupled with tandem MS (MS/MS) analysis at the W.R. Wiley Environmental Molecular Sciences Laboratory identified greater than 4400 cellular proteins along with 7 HCV proteins with high confidence after the application of conservative search criteria coupled with LC elution time information (Figure 1). A semi-quantitative comparison of total peptide identifications was subsequently used as a first pass means to detect changes in protein abundance associated with HCV RNA replication. Preliminary comparative analysis of protein abundance in the presence



4,474 total proteins identified

7 of 10 HCV proteins were identified, all within the microsomal fraction of the (+) HCV sample

Viral Proteins	Peptide #
HCV core protein	2
HCV glycoprotein E1	0
HCV glycoprotein E2	3
HCV p7 protein	0
HCV NS2 proteinase	5
HCV NS3 proteinase/helicase	10
HCV NS3/4A proteinase cofactor	0
HCV NS4B protein	3
HCV NS5A phosphoprotein	6
HCV NS5B RNA dependant RNA polymerase	14

	Total	Total (-) HCV	Overlap (+)&(-) HCV	Total (+) HCV
Total Peptides	24,921	16,967	7,660	15,614
Viral Peptides	48	0	0	43
Viral Proteins	7	0	0	7
Total Proteins	4,474	3,570	2,413	3,317

Figure 1. Schematic representation of sample preparation, separation, and analysis of the Huh-7.5 replicon model system with tables of results showing the number of identified peptides for both host cellular and viral proteins.

and absence of the full-length replicon revealed potential up-regulated proteins involved in many cellular stress-related processes and down-regulated proteins involved in multiple metabolic pathways. A majority of these observations have been previously reported in the literature for HCV liver infections, while others appear to be novel discoveries.

The use of a full-length HCV replicon offers the advantage of allowing investigation of the potential influence of both structural and nonstructural proteins on the host cell, thus providing a more comprehensive view of the interactions among multiple pathways responding to viral infection. These current findings provide additional evidence supporting a generally emerging theme whereby disruptions in metabolic pathways induce a state of cellular stress that contributes to the pathogenic effects of HCV. The results discussed here clearly demonstrate the potential for proteomic studies of the HCV replicon system to assist in the determination of proteins/pathways affected by HCV infection. However, this model is somewhat limited in its inability to support production of infectious viral particles, which precludes analysis of the complete viral life cycle *in vitro*. Moreover, this system does not provide a model for the study of human liver biology and disease. We have therefore begun to extend our analyses to serial liver biopsy specimens obtained from patients with recurrent HCV after liver transplantation. Serial liver biopsies represent a true system of HCV infection and liver disease. Moreover, the liver transplant model offers the unique opportunity to prospectively examine early viral host interactions and to determine which of these factors show critical associations with severe liver injury and early progression to cirrhosis.

Preliminary studies using the accurate mass and time (AMT) tag approach for high-throughput, high-sensitivity Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) analysis identified greater than 1500 proteins from less than 1μg of protein lysate prepared from liver biopsy tissue (Figure 2). Previous studies of the liver proteome involving one-dimensional or two-dimensional polyacrylamide electrophoresis for protein

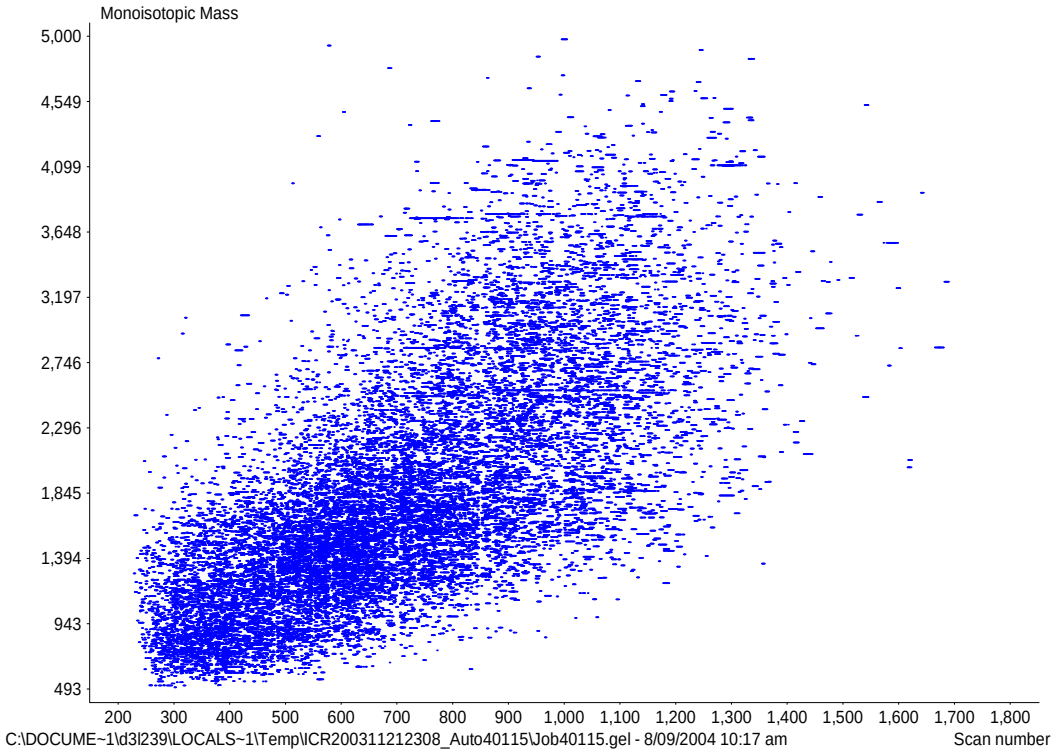


Figure 2. Two-dimensional display of a single FTICR-MS analysis of less than 1 μg of liver biopsy tissue, within which more than 7000 peptides were identified.

separation coupled with other detection methods typically need mg amounts of protein lysate to identify fewer than 300 proteins. The combination of high-quality, liquid-phase separations and unique, high-resolution FTICR instrumentation provides even greater sensitivity and dynamic range now allowing for identification of thousands of proteins from μg rather than mg amounts of starting material. This robust and ultra-sensitive 'nano-proteomic' technology was critical to our success in analyzing the highly complex proteome of small liver biopsies. The low protein yields (often less than 50 μg total protein) associated with these small liver biopsy specimens precluded previous attempts to detect a broad abundance range of proteins using more conventional methods (e.g., isotope-coded affinity tagging + LC-MS/MS). These studies lay the foundation for current efforts aimed at employing the AMT tag strategy together with stable isotope labeling methods for quantitation of differential protein abundance during the clinical course of HCV infection. The data generated from our *in vitro* (HCV replicon) and *in vivo* (liver transplant model) proteomic studies will support ongoing efforts to develop and visualize computational models that integrate mRNA and protein abundance data with the host-virus network scaffold to generate predictive models of host-virus network states. This information is expected to facilitate the identification of biomarkers for HCV infection and disease progression as well as novel targets for future antiviral treatment.

The power of FTICR-MS represents a significant advancement in clinical proteomics efforts and offers the unique opportunity to begin investigating the clinical significance of protein

abundance changes associated with HCV infection. The key capability that these types of analyses may make possible is the analysis of the complete proteome in a single experiment. This includes

- identification and quantitation of all of the proteins present with high confidence
- completion of the analysis in a clinically relevant time period of approximately two hours
- use of minimal amounts of material so that samples can be taken from patient with as little trauma as possible.

The full realization of these capabilities will take further improvements in the technology, but these studies clearly demonstrate the potential of the methodology.

We are particularly interested in further characterizing the mechanisms of HCV-induced alterations in metabolic pathways that contribute to the development of cellular stress and necrosis. The apparent link with progression of liver injury and fibrosis makes these cellular metabolic pathways attractive targets for discovering new anti-fibrotic drugs.

Proteomic Characterization of Cerebrospinal Fluid (CSF) by High-Resolution LC-MS/MS

JM Jacobs,[(a)] DG Camp,[(a)] and S Warren[(b)]

(a) Pacific Northwest National Laboratory, Richland, Washington

(b) Massachusetts General Hospital East and Harvard University School of Medicine, Charlestown, Massachusetts

Human cerebral spinal fluid is a simpler mixture than blood, and studying it directly should provide a shorter path to diagnosing illnesses of the brain or spinal cord. In this study, the proteome of human cerebral spinal fluid is characterized for the first time and will be used to identify specific biomarkers for the clinical diagnosis of these illnesses.

In this project, a researcher in the Infectious Disease Unit of Massachusetts General Hospital and scientists at Pacific Northwest National Laboratory have obtained the first-ever results from the proteome analysis of human cerebral spinal fluid (CSF) by mass spectrometry (MS). CSF is routinely sampled in clinical medicine to help determine and monitor the presence of illness in the brain or spinal cord. Clinicians usually measure the total protein concentration for this purpose. This approach provides no detailed information about the specific illness that might be causing a change in the protein concentration, and a more detailed look at the proteins present would be of little use because very little is known about the protein composition, or proteome, of CSF in either normal or disease states. Much of our current knowledge of CSF proteins is based on two-dimensional gel analysis techniques, which are severely limited in sensitivity, specificity, and throughput.

To better understand the protein composition of CSF and to create a baseline of information for comparative analysis of disease states, this project will create as detailed a list of CSF proteins as possible using the world's most advanced proteomics capabilities, located at the W.R. Wiley Environmental Molecular Sciences Laboratory (EMSL). These capabilities include ultra-high-pressure liquid chromatography (LC) systems used to separate peptides characteristic of the proteins in a sample and characterize them chemically by their elution time, extensive tandem mass spectrometry (MS/MS) capabilities used to create comprehensive databases for a biological system, and Fourier transform ion cyclotron resonance (FTICR) mass spectrometers used for fast analysis of samples against the database.

The overall approach is divided into two phases. Phase 1 will determine, to the greatest extent possible, the content and distribution of proteins in 'normal' human CSF. This information will be used for the creation of a mass and time tag database for future quantitative LC-FTICR studies. Phase 2 will evaluate how the content and abundance of proteins in CSF are altered in specific diseases of the brain and spinal cord.

Phase 1 of the project has proceeded rapidly with a series of experiments to optimize the sample processing of CSF for the baseline studies. We have received two CSF clinical samples that we submitted to multiple sample preparation techniques, after which we performed MS/MS analysis to assess each method for the detection of peptides from each sample. This is the first such analysis of CSF clinical samples by LC-MS. Detection of CSF

proteins presents many challenges (e.g., high glycosylation, the overall low-protein content of CSF, small sample sizes), which are being overcome. It is precisely these challenges that require the high-sensitivity MS instrumentation in EMSL.

Phase 1 efforts are continuing, and shortly we will begin to fractionate the CSF sample to increase the dynamic range of detection and thereby maximize the protein coverage. With the completion of these analyses, a normal baseline of proteins will be established that will be the most comprehensive protein characterization of human CSF ever created. This will provide a unique resource for future studies. For Phase 2, a well-defined disease of the brain or spinal cord, most likely an infectious disease, will be selected. Samples of CSF will be analyzed to determine the quantity of both human host proteins and microbial proteins for comparison against normal CSF. These efforts would represent the first step in developing specific biomarkers based on proteins that are present in CSF only or predominantly as the result of a microorganism-specific infection. Such information could be vital for bedside clinical diagnoses and would be a significant improvement over present procedures.

A Proteomic Approach to Characterize Protein Shedding

M Ahram,[(a)] JN Adkins,[(a)] DL Auberry,[(a)] DS Wunschel,[(a)] and DL Springer[(a)]
(a) Pacific Northwest National Laboratory, Richland, Washington

Knowing what could be present in the sample and when it might be present is a requirement for using body-fluid-based samples to diagnose an individual's state of health. New mass spectrometry technology permits the detection and identification of proteins from such samples. This study is an important step in characterizing the proteins that cells shed into the body, specifically from cultured human mammary epithelial cells. Forty-five different proteins were detected, and it was demonstrated that their abundance could be used to indicate the biological state of those cells.

Introduction

Cells can interact with each other and with their environment by excreting and absorbing proteins or parts of proteins from the surrounding medium. One process by which they do this, referred to as shedding, involves the cleavage (proteolysis) of the portion (ectodomain) of transmembrane proteins that extends into the surrounding media. Protein turnover plays an important physiological role in disease. Membrane proteins of different topologies are released from the cell surface by shedding proteases (sheddases). Substrates for the sheddases include receptors, cell adhesion proteins, and structural proteins. The main family of proteases involved in shedding is the A desintegrin and metalloprotease (ADAM) family; a notable example is tumor necrosis factor-alpha (TNF-α) converting enzyme (TACE), which was the first member of the ADAM family identified as a sheddase.

Protein shedding can be up-regulated by certain agonists such as phorbol esters and specifically 4β-phorbol 12-myristate 13-acetate (PMA). Because of the biological significance of protein shedding, there is a need to identify shed proteins. The use of proteomic methods, in particular mass spectrometry (MS), has shown great promise in protein discovery. In this project, we undertook the first proteomics study by MS of proteins shed from serum-free media of human mammary epithelial cells grown in culture after induction of shedding by the phorbol ester, PMA.

Results

Sample Preparation and Mass Spectrometry Analysis

Different preparation methods for media from PMA-treated and control cells were performed to determine the best method of sample processing and to enrich the samples for a wide variety of proteins. The first method involved the attachment of biotin (biotinylation) to cell surface molecules. Almost universal biotinylation of surface proteins was achieved. PMA then was administered to the cell to induce the shedding of biotinylated proteins. Then the bitinylated proteins present in the media were isolated using a streptavidin-affinity column. The second method involved the induction of shedding with PMA without modification of surface proteins. Proteins with attached sugars (glycans) then were isolated using lectin-affinity chromatography. The samples produced by the second method were analyzed both as is and after removal of the glycans. A third sample preparation involved

collection and analysis of the total protein populations present in the media following treatment with PMA and concentration by filtration with a low molecular weight cut-off filter and precipitation using trichloroacetic acid (TCA) of low molecular weight peptides that were not removed by the filter. The total protein quantities resulting from the control and PMA treatments for the three sample preparation methods were not significantly different.

The biotinylated, glycosylated, and crude protein fractions were then digested into peptides with the enzyme trypsin, which cleaves the proteins at the amino acid residues lysine and arginine. The samples were analyzed using reversed-phase capillary liquid chromatography directly coupled to an ion trap mass spectrometer by electrospray ionization (ESI). The mass spectrometer identified the molecular weight of peptides present in the sample and automatically selected the most intense ions for further fragmentation and subsequent analysis of these fragments. These tandem mass spectrometry (MS/MS) spectra were analyzed by SEQUEST, and protein abundances were estimated.

MS data from 10 sample preparations were evaluated including: three that represented biotinylated proteins, three contained purified glycosylated protein, and four represented total populations of proteins. The samples were analyzed six times except in the cases of biotinylated and lectin-purified samples, most of which were evaluated twice because of limited sample availability. SEQUEST analyses of MS/MS spectra identified 908 peptides representing 201 proteins (a complete listing of these proteins, the preparation method, and status of their membranes association is presented in Table 1). Because some proteins were observed in more than one fraction, the total number of proteins identified is less than the sum number of proteins shown in Table 1. It is also worth noting that deglycosylation of media samples did not result in improvement in the number of protein identifications.

Table 1. Quantity of the 221 proteins identified in the three preparations according to their extracellular membrane localization.

	Biotinylated	Glycosylated	Crude	Total
Membrane	7	4	37	48
Non-membrane	15	12	137	164
Unknown	2	0	7	9

Two proteins, syndecan-4 and hepatoma-derived growth factor (HDGF), were identified as molecules potentially shed into the media. Interestingly, we also identified proteins, including glucose-regulated protein 78 kDa (GRP78), whose abundances decreased subsequent to PMA treatment. Immunoblotting methods validated our proteomic findings for three proteins, syndecan-4, GRP78, and CD44. These results indicate the feasibility of identifying shed proteins using a global proteomic MS-based approach.

Identification of shed proteins

To identify proteins that are likely derived from the shedding process, we evaluated the extracellular nature of the identified proteins by predicting the presence of membrane helices. Proteins that did not contain transmembrane domains but were known to associate to the cell surface in a non-transmembrane domain manner, a receptor, or an adapter protein also were included. Secreted proteins known to associate with membrane proteins also were

considered. Using these criteria, we identified 45 membrane-associated proteins (22.3% of the total proteins identified) from the three fractions. Of these, 22 proteins contained transmembrane domains. Moreover, 18 of the 22 proteins were identified by one or more extracellular peptides. Thus, these 18 proteins are strong candidates as shed proteins and likely enter the extracellular media through the process of regulated proteolysis. We also identified 23 proteins that associate to the cell surface indirectly and may be released into the media via shedding of membrane-integral proteins. Eight proteins of unknown localization were observed including transketolase and heparin-binding, growth-factor-binding protein.

Most of the 45 membrane-associated proteins originated from the crude preparation (37 proteins); whereas, seven of the proteins were detected in the biotinylated fraction, and four were in the glycosylated fraction. These results indicate that even though the crude preparation contained a high proportion of non-membrane proteins, this preparation generated the largest number of membrane-associated proteins and, hence, likely shed.

Protein abundance

Previously, we observed a correlation between the number of peptides detected by MS and protein abundance (Adkins et al. 2002). We used this method to estimate protein abundance in the media following either control or PMA treatment. The 201 proteins were evaluated using an additional set of rules to identify changes in protein abundance as a result of PMA treatment. The principal rule provided that because we are attempting to identify shed proteins, the protein must associate with the cell surface either directly or indirectly or have an unknown localization. Upon applying these rules, the abundances of six proteins were defined as altered by PMA treatment. Two of these proteins, syndecan-4 and HDGF, were found in higher abundances in media of PMA-treated cells. On the other hand, we concluded that the abundances of four proteins decreased including enolase I, GRP78, lactate dehydrogenase A, and transketolase.

We validated protein abundance estimates using immunoblotting methods. The level of syndecan-4 protein increased in the PMA-treated samples by a factor of 2.3 relative to controls upon quantification of band intensities, whereas the abundance of GRP78 decreased by a factor of 2.0. On the other hand, the level of CD44E was unaltered. These results were in good agreement with the abundance estimates using the peptide count method and in support of the additional rules appended to the SEQUEST criteria.

In this study, we used different methods to identify shed proteins by MS and found the most protein identifications with the total protein preparation without chemical modification (e.g., biotinylation) or affinity separation (e.g., lectin-chromatography). We identified 18 proteins that are strong candidates as shed proteins in that they contained transmembrane domains and were identified by peptides derived from the extracellular portion of the protein. We also identified 23 proteins that associate to the plasma membrane via binding to integral membrane proteins and could be released as a result of shedding of these proteins. Upon estimating the abundance of proteins released in the media by counting the number of peptides detected, we were able to identify six proteins whose abundances were altered by PMA treatment. The abundances of three proteins were validated by traditional immunodetection methods. Our method provides protein abundance estimates that can be used without the additional workup and expense associated with isotope-coded affinity tagging

(ICAT) (Gygi et al. 1999). However, in the case where measurements of smaller changes in protein abundances are required, the ICAT method may provide additional precision. In addition, it is critical to couple identification of shed proteins using MS approaches with immunological techniques, when possible, to provide high-confidence information.

Acknowledgements

This research was supported by the Office of Science (Office of Biological and Environmental Research) U.S. Department of Energy Low Dose Program.

References

Adkins JN, SM Varnum, KJ Auberry, RJ Moore, NH Angell, RD Smith, DL Springer, and JG Pounds. 2002. "Toward a Human Blood Serum Proteome: Analysis by Multidimensional Separation Coupled With Mass Spectrometry." *Molecular and Cellular Proteomics* 1(12):947-955.

Gygi SP, B Rist, SA Gerber, F Turecek, MH Gelb, and R Aebersold. 1999. "Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags." *Nature Biotechnology* 17(10):994-999.

User Projects

Gas Phase Reactions of Iron-Silicon Clusters

Y Fu, L Wang
Washington State University, Tri-Cities, Richland, Washington

Identification of Post-Translational Modifications and Protein Complexes Under Conditions of Oxidative Stress

CA Sacksteder, TC Squier
Pacific Northwest National Laboratory, Richland, Washington

Proteomic Analysis of the Human Mammary Epithelial Cell Mitogenic Response

BD Thrall
Pacific Northwest National Laboratory, Richland, Washington

Proteomics of *Shewanella oneidensis* MR-1 Subjected to Solar Ultraviolet Radiation Stress

JM Tiedje
Michigan State University, East Lansing, Michigan

Identification of Protein Components of Vaccinia Virus Particles

B Moss
National Institutes of Health, Bethesda, Maryland

Identifying Targets for Therapeutic Interventions Using Proteomic Technology

JN Adkins, KD Rodland
Pacific Northwest National Laboratory, Portland, Oregon
HM Mottaz
W.R. Wiley Environmental Molecular Sciences Laboratory, Richland, Washington
F Heffron, S Wong
Oregon Health Sciences University/Oregon Graduate Institute, Beaverton, Oregon

Proteomic Characterization of Cerebrospinal Fluid by High-Resolution Liquid Chromatography Tandem Mass Spectrometry

S Warren
Massachusetts General Hospital East and Harvard University School of Medicine, Charlestown, Massachusetts
JM Jacobs, DG Camp
Pacific Northwest National Laboratory, Richland, Washington

Proteomic Analysis of the Cyanobacterium *Synechocystis* Species PCC 6803

JM Jacobs, DG Camp
Pacific Northwest National Laboratory, Richland, Washington
HB Pakrasi, N Keren
Washington University in St. Louis, St. Louis, Missouri

Characterization of the Tumor Cell Lamellipodia Phosphoproteome
RM Klemke
Scripps Research Institute, La Jolla, California
JM Jacobs, DG Camp
Pacific Northwest National Laboratory, Richland, Washington

Unraveling the Molecular Biology of Host-Pathogen Interactions
L Shi
Pacific Northwest National Laboratory, Richland, Washington

Peptide Observation Model
LJ Kangas
Pacific Northwest National Laboratory, Richland, Washington
LS Wang
Washington State University Tri-Cities, Richland, Washington

Characterization of Neuropeptides Located in Neurosecretory Tissue of Cancer Borealis and Cancer Productus
KK Kutz
University of Wisconsin, Madison, Wisconsin

Whole Genome Cloning and Expression of *Treponema pallidum* Open Reading Frames
TG Palzkill
Baylor College of Medicine, Houston, Texas

Proteomic Analysis of Wild-Type and Soil-Colonization-Deficient Strains of *Pseudomonas fluorescens*
SB Levy, I Lopez-Hernadez, EA Robleto
Tufts University School of Medicine, Boston, Massachusetts
MS Lipton
Pacific Northwest National Laboratory, Richland, Washington

Comparative Display of *Deinococcus radiodurans* After Exposure to Ionizing Radiation
MJ Daly
Oak Ridge National Laboratory, Bethesda, Maryland

Hollow Helices as Folding Nanotubes with Tunable Cavity Size
B Gong
State University of New York at Buffalo, Buffalo, New York

Determination of the Proteome of Drosophila Lipid Droplets
SP Gross, Je Martinez
University of California, Irvine, Irvine, California

ICAT Labeling of *Pseudomonas aeruginosa* Membrane Subproteome
W Xiao, RW Davis
Stanford University, Stanford, California

Mechanism of Action of G Protein-Coupled Receptors Studied by Liquid Chromatography-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry
EA Dratz, P Kraft
Montana State University, Bozeman, Montana

Identification of Heat Shock Proteins in the Radioresistant Bacterium *Deinococcus radiodurans*
AK Schmid
University of Washington, Seattle, Washington

Identification of Proteins Differentially Expressed in Response to Infrared Radiation
MA Coleman
Lawrence Livermore National Laboratory, Livermore, California

Identification and Relative Expression of Membrane Proteins in Breast Cancer Cell Lines
AJ Patwardhan, K Williams, M Pallavicini
University of California, San Francisco, San Francisco, California
EF Strittmatter, R Smith, L Pasa-Tolic, DG Camp
Pacific Northwest National Laboratory, Richland, Washington

Identification of a Novel Biomolecule in Deep-Sea Clams
PH Yancey
Whitman College, Walla Walla, Washington

Posttranslational Modifications of Chicken Brain Myelin Basic Protein Charge Isomers
R Zand
University of Michigan, Ann Arbor, Michigan

High Throughput Fourier Transform Ion Cyclotron Resonance Strategies for the Molecular Analysis of Cancer
DG Camp
Pacific Northwest National Laboratory, Richland, Washington

Determination of the C-Terminus of p110 Soluble Epidermal Growth Factor Receptor
RC Zangar
Pacific Northwest National Laboratory, Richland, Washington

Proteomic Characterization of *Yersinia pestis*
MS Lipton, SL McCutchen-Maloney
Pacific Northwest National Laboratory, Richland, Washington

Biomolecular Mechanisms for Microbe-Iron (III) Oxide Interactions in *Geobacter* Species
SE Childers
University of Idaho, Moscow, Idaho
TJ Beveridge
University of Guelph, Guelph, Ontario, Canada
MS Lipton
Pacific Northwest National Laboratory, Richland, Washington
DR Lovley
University of Massachusetts, Amherst, Amherst, Massachusetts

Proteomics of Morphology Determination in a Fungus
EA Panisko, JK Magnuson, LL Lasure
Pacific Northwest National Laboratory, Richland, Washington

Regulated Proteolysis and Bystander Effects in Radiation
DL Springer
Pacific Northwest National Laboratory, Richland, Washington

Global Measurments of Phosphopeptides in Human Mammary Epithelial Cells
RC Zangar
Pacific Northwest National Laboratory, Richland, Washington

Preliminary Work on the Proteomes of Brains and Dissected Brains Obtained from Control Mice and Treated Mice Simulating Parkinsons Disease
D Smith
University of California, Berkeley, Berkeley, California
DG Camp, DM Sforza
Pacific Northwest National Laboratory, Richland, Washington

Molecular Dysfunction in Alcoholic Liver Disease
A Banerjee
Wayne State University, Detroit, Michigan

Engineering Device Surfaces that Instruct Cell Behavior
PS Stayton, A Golden
University of Washington, Seattle, Washington

Abundance of Protein Components in Photosystem II Protein Complex Purified from Mutant Cyanobacterial Cells Lacking Individual Subunit Protein
HB Pakrasi, N Keren
Washington University, St. Louis, Missouri

Using Mass Spectrometry to Assist in Identifying Genes for Enzymes of Known Activity, Unknown Open Reading Frame, and Small Abundance
DH Kohl
Washington University, St. Louis, Missouri

Determination of the Peptide-Binding Sites on the Bacterial Chaperone Protein SecB
LL Randall
University of Missouri, Columbia, Columbia, Missouri
VF Smith
United States Naval Academy, Annapolis, Maryland

Mass Spectrometric Analysis of Eukaryotic Proteosome
RA Maxwell
Pacific Northwest National Laboratory, Richland, Washington

Mass Spectrometry Analysis of Nucleotides/Nucleosides
K McAteer
Washington State University Tri-Cities, Richland, Washington
NG Isern
W.R. Wiley Environmental Molecular Sciences Laboratory, Richland, Washington

Cellular Response to Hepatitis C Virus Infection: Global Quantitative Proteome Accurate Mass and Time Tag Measurements of Cellular Protein Expression
MG Katze, DL Diamond, EY Chan
University of Washington, Seattle, Washington

Protemic Analysis of Naturally Ocurring Methane Oxidizing Archaeal Communities
SJ Hallam, EF DeLong, PR Girguis
Monterey Bay Aquarium Research Institute, Moss Landing, California

Identification of the SecB Residue that is Labeled by Methanethiosulfonate
LL Randall, JM Crane
University of Missouri, Columbia, Columbia, Missouri

A Proteomic Approach to the Study of *Neurospora crassa* Ribosomal Dynamics
SE Baker, EA Panisko
Pacific Northwest National Laboratory, Richland, Washington
M Sachs
Oregon Health Sciences University/Oregon Graduate Institute, Beaverton, Oregon

Proteomic Analysis of *Fusarium graminearum* Sexual Development
SE Baker, EA Panisko
Pacific Northwest National Laboratory, Richland, Washington
BG Turgeon
Cornell University, Ithaca, New York

Proteomics of Membrane Protein Complexes
DJ Bigelow, CA Sacksteder
Pacific Northwest National Laboratory, Richland, Washington

Using the Approaches of Mass Spectrometry-Based Phosphoproteomics and Small Interfering RNA (siRNA)-Induced Gene Silencing to Identify New Insulin Receptor Signaling Components in Cultured Adipocytes
ZY Jiang
University of Massachusetts Medical School, Worcester, Massachusetts

Proteome Analysis of *Deinococcus radiodurans* Exposed to Chronic Irradiation
MJ Daly
Oak Ridge National Laboratory, Bethesda, Maryland

Genomes-to-Life Protein Complex Isolation Capability
TC Squier, BS Hooker, HS Wiley, RW Siegel, V Kery, MU Mayer-Cumblidge, L Shi, JN Adkins, G Michaels
Pacific Northwest National Laboratory, Richland, Washington

Proteomics of Filamentous Fungi
JK Magnuson, SE Baker, EA Panisko, LL Lasure, Z Dai
Pacific Northwest National Laboratory, Richland, Washington

Sustrate Identification for Ser/Th Protein Phosphatase
SS Rossie
Purdue University, West Lafayette, Indiana

Identification of Functional Pathways Associated with Clinical Tamoxifen Resistance in Breast Cancer by Advanced Mass Spectrometry
A Umar, T Luider
Erasmus University Medical Center, Rotterdam, Rotterdam, Netherlands

Characterization of Gold Nanocluster Materials
H Zhang, L Wang, HM Lee, X Huang
Washington State University Tri-Cities, Richland, Washington

Cellular Response to Human Immunodeficiency Virus Type 1 Infection: Global Quantitative Proteome Accurate Mass and Time Tag Measurements of Cellular Protein Expression
EY Chan, MG Katze, DL Diamond
University of Washington, Seattle, Washington

Proteomic Analysis of *Shewanella oneidensis* Biofilms
AM Spormann, R Saville
Stanford University, Stanford, California

Membrane Proteins of Placental Parasites
M Fried
Seattle Biomedical Research Institute, Seattle, Washington

Combined Transcriptome and Proteome Analysis of *Shewanella oneidensis* MR-1 etrA and arcA Mutant Strains
YA Gorby
Pacific Northwest National Laboratory, Richland, Washington
SD Brown, J Zhou
Oak Ridge National Laboratory, Oak Ridge, Tennessee

Proteomics of Mitochondria
C Scharfe, P Oefner
Stanford University, Stanford, California

Photoinhibition of Photosynthesis
H Hou
Gonzaga University, Spokane, Washington

Implementation of Systems Toxicology for an Animal Emphysema Model
JG Pounds
Pacific Northwest National Laboratory, Richland, Washington

Proteomic Analysis of Arthrobacter
A Konopka
Purdue University, West Lafayette, Indiana

Proteomic Studies of Growth and Metabolism in *Corynebacterium glutamicum* R
FB Metting, SE Baker
Pacific Northwest National Laboratory, Richland, Washington
M Inui, H Yukawa
Research Institute of Innovative Technology for the Earth, Kizu, Japan

Staff

Harold R. Udseth, Technical Lead
(509) 376-3698, harold.udseth@pnl.gov

Kimberly S. Korenkiewicz, Administrator
(509) 373-0765, kimberly.korenkiewicz@pnl.gov

Dave J. Anderson, Scientist
(509) 376-7898, david.anderson@pnl.gov

Kim K. Hixson, Scientist
(509) 373-6170, kim.hixson@pnl.gov

Suzana Martinović, Scientist
(509) 376-3709, suzana.martinovic@pnl.gov

Ronald J. Moore, Technologist
(509) 376-2721, ronald.moore@pnl.gov

Heather M. Mottaz, Scientist
(509) 376-4869, heather.mottaz@pnl.gov

Rui Zhang, Scientist
(509) 376-3200, rui.zhang@pnl.gov

Rui Zhao, Scientist
(509) 376-7332, rui.zhao@pnl.gov

We would also like to acknowledge the contributions of David G. Camp, Carrie D. Goddard, Marina Gritshenko, Mary S. Lipton, Christophe D. Masselon, and Ljiljana Pasa-Tolic.