PRESENTER ABSTRACTS

Genotypic Variability and the Quantitative Proteotype

Ruedi Aebersold, Department of Biology, Institute of Molecular Systems Biology, ETH Zurich and Faculty of Science, University of Zurich

The question how genetic variability is translated into phenotypes is fundamental in biology and medicine. Powerful genomic technologies now determine genetic variability at a genomic level and at unprecedented speed, accuracy and (low) cost. To date the effects of genomic variability on the expressed information of the cell has been mainly studied by transcript profiling.

In this presentation we will discuss emerging computational and quantitative proteomic technologies to relate genotypic variation to the proteome, specifically to protein abundance, protein modification and the organization of proteins into functional modules. Proteomic data to support such correlations need to be quantitatively accurate, highly reproducible across multiple measurements and samples and generated at high throughput. The SWATH-MS technology we recently developed is capable of generating datasets with these properties for thousands of proteins from complex sample mixtures.

We will discuss the principles and limitations of SWATH-MS, discuss its computational challenged for data analysis and demonstrate with selected applications, using genetic reference strain compendia, the ability of the technology to determine the effect of genetic variability on the quantitative proteome at the level of protein abundance and post translational modification thus functionally connecting the genome to proteome and phenotype.

Phosphorylation-regulated Protein Dynamics in Kinase Regulation and Implications for Inhibitor Design: The Case of ERK2

Natalie G. Ahn, Department of Chemistry and Biochemistry, BioFrontiers Institute, University of Colorado, Boulder, CO, USA

The MAP kinases, extracellular-regulated protein kinases 1 & 2 (ERK1/2), are important drug targets for cancers caused by oncogenic mutations in RAS and B-RAF. Preclinical studies show that cells from metastatic cancers with acquired resistance to RAF and MKK inhibitors can be effectively killed using small molecule inhibitors of ERK, two of which are in clinical trials. ERK1/2 are activated by dual phosphorylation at Thr and Tyr residues, both which are catalyzed by MKK1/2. We examined ERK2 activation using hydrogen exchange mass spectrometry (HX-MS), which can report localized conformational mobility within folded proteins, where exchange predominantly occurs through low energy fluctuations in structure, allowing transient solvent exposure. Changes in conformational mobility may impact protein function, even when structural changes are unobservable. Phosphorylation of ERK2 leads to changes in local HX, suggesting that kinase activation modulates protein motions. This was corroborated by NMR relaxation dispersion experiments, which revealed substantial changes in the dynamics of the enzyme upon phosphorylation and activation. Preliminary data suggest a model in which activation of ERK2 leads to global exchange between two conformational states, where phosphorylation-regulated dynamics are coupled to steps in catalytic turnover. We further investigated high affinity ERK inhibitors which belong to families that have been shown to be effective towards cells with acquired resistance. We find that one inhibitor shifts the equilibrium completely towards one conformational state, while the other inhibitor shifts the equilibrium completely to the opposite state. Therefore, ERK inhibitors have properties of conformation selection. We propose that ERK2 is regulated at the level of protein motions that facilitate the catalytic cycle, and that this dynamic behavior may be exploited to improve properties of ERK inhibitors.

An Edgotype of Chemoresistant Cells Reveals the Shaping of Epigenetic Regulatory Components

James E. Bruce, School of Medicine, University of Washington, Seattle, WA

Chemoresistance is a common mode of therapy failure for many cancers. Tumors develop resistance to chemotherapeutics through a variety of mechanisms, with proteins most often serving pivotal roles. Changes in protein conformation, localization and interactions, often driven by post-translational modifications, affect cellular response to environmental conditions contributing to the development of new phenotypes. The ability to decipher how protein interaction networks adapt and yield new function or alter phenotype has been limited by the inability to visualize large-scale changes in protein structural features and protein-protein interactions. Recent developments in chemical cross-linking and mass spectrometry technologies are beginning to allow visualization of protein structural features and interactions in live cells, including those driven by post-translational modifications. Identification of cross-linked peptide sequences reveals protein-protein interactions and structural features of proteins and complexes. Our current efforts with multidrug resistant human carcinoma cells demonstrate that quantitative interactome information can be acquired with this approach. These efforts enabled quantitative analysis of the largest protein interaction network generated to date from *in vivo* chemical cross-linking. Detection and quantitation of 1391 cross-linked peptide pairs enables so-called 'edgotype' analysis in cancer cells with acquired chemoresistance. This presentation will describe our quantitative approach and demonstrate its application to sensitive and resistant cancer cells to reveal changes in protein interactions and structures that correlate with a chemoresistant phenotype.

Fast and Comprehensive Proteome Analysis

Joshua Coon, Department of Chemistry, University of Wisconsin-Madison, Madison, WI

Modern biological and medicinal researchers increasingly rely on technologies for rapid and large-scale the comparative analysis of proteomes. Here we describe the first is the identification of ~4,000 yeast proteins, a comprehensive yeast proteome, from just over an hour of analysis time. We achieved this expedited proteome characterization through improved sample preparation, chromatographic separations, and by using a new Orbitrap hybrid mass spectrometer equipped with a mass filter, a collision cell, a high-field Orbitrap analyzer, and, finally, a dual cell linear ion trap analyzer (Q-OT-qIT, Orbitrap Fusion). Coupling this technology with a multi-enzyme digestion approach and multiple MS/MS dissociation techniques (i.e., ETD, HCD, CAD), we also report the deepest coverage achieved to date of the yeast proteome. In this work we detect 5,033 yeast proteins (detection of 3 dubious ORFs, two of which had very high sequence coverage) and had a median sequence coverage of > 80%. Specifically, we recorded mass spectral evidence for 8 of every 10 amino acids in the yeast proteome.

Dynamic Protein Modification in Metabolism and the Epigenome

John Denu, Biomolecular Chemistry, Epigenetics, Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI

Dynamic chemical modifications to proteins can function as a signaling mechanism, communicating extracellular events and intracellular metabolic dynamics. Addition of many post-translational modifications (PTMs) such as methylation, acetylation and phosphorylation, and removal of these groups involve enzymes that utilize key metabolic intermediates as co-substrates. This intricate dependence on these metabolites suggests that many protein modifications are regulated by the normal ebb and flow of metabolite levels, in response to diet, lifestyle and other environmental factors. Dysregulation of these mechanisms could be causal or a manifestation of genetic and epigenetic diseases. Having quantitative proteomic tools to address these questions is key to providing the necessary mechanistic understanding. The Denu group, in collaboration with a number of other laboratories, is developing and utilizing proteomic tools to quantify dynamic changes to the acetylproteome and the histone PTM code. In this talk, Denu will present recent studies that demonstrate the role NAD^+-dependent protein acetylation in multiple mammalian tissues and its impact on inter-tissue metabolism. The importance of quantifying site-specific acetylation stoichiometry will be presented. Also, Denu will discuss how quantitative analysis of ~80 histone PTM states provides a unique window into how alterations from metabolism, diet or environment affect the dynamics of epigenetic information written onto histones.

Top-Down Proteomics for Comprehensive Analysis of Post-translational Modifications: Application to Heart Failure

Ying Ge, Department of Cell and Regenerative Biology, Department of Chemistry, and Human Proteomics Program, University of Wisconsin-Madison

Post-translational modifications (PTMs) modulate protein activity, stability, and function, playing essential roles in many critical cell signaling events in healthy and disease states. Dysregulation of PTMs have been implicated in a spectrum of human diseases. Hence, a comprehensive analysis of PTMs is imperative for a better understanding of human diseases but presents tremendous challenges due to the complexity and heterogeneity of PTMs.

Top-down mass spectrometry (MS)-based proteomics, based on the analysis of intact proteins, is arguably the most powerful technology to comprehensively characterize PTMs together with genetic variations. We have shown that top-down MS has unique advantages for unraveling complex protein PTMs, mapping modification sites with full sequence coverage, discovering unexpected modifications, quantifying multiple PTMs and monitoring PTMs changes during the disease progression. However, top-down proteomics still faces significant challenges in terms of protein solubility, protein separation, detection of low-abundance and high-mass proteins, and under-developed software packages. Recently, we are employing a multi-pronged approach to address these challenges in a comprehensive manner by developing new MS-compatible surfactants for protein solubilization, new strategies for multi-dimensional chromatography separation of proteins, novel nanomaterials for enrichment of low-abundance proteins, and a new comprehensive software package for top-down proteomics. Meanwhile, we have employed the top-down proteomics platform to study heart diseases. Importantly, we have successfully linked altered cardiac protein PTMs to contractile dysfunction in heart failure using both animal models and human clinical samples**.** In this presentation I will discuss our recent technology developments in top-down proteomics for comprehensive characterization of PTMs and their application to understand heart failure.

O-GlcNAcylation Serves as a Nutrient Sensor to Regulate Transcription, Signaling and Mitochondrial Function, and is a Major Molecular Basis of Glucose Toxicity in Diabetes

Gerald W. Hart, Department of Biological Chemistry, Johns Hopkins Univeristy School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205

O-GlcNAcylation is the cycling of N-acetylglucosamine residues on serine or threonine moieties mostly on nuclear and cytoplasmic proteins. O-GlcNAcylation serves as a sensitive nutrient sensor since its cycling and protein specificity are highly dependent upon the concentrations of UDP-GlcNAc, which in turn is linked to the major metabolic cellular pathways, including glucose, amino acid, fatty acid, nucleotide and energy metabolism. In addition to modifying nearly all proteins involved in transcription, including histones, RNA polymerase, DNA methyltransferases, and both basal and specific transcription factors, O-GlcNAc cycles on well over half of protein kinases. Crosstalk with protein phosphorylation is extensive, and thus far, kinases modified by O-GlcNAcylation are indeed regulated by the sugar. As a result of O-GlcNAc's extensive roles in both transcription and signaling, the sugar plays a key role in glucose toxicity in diabetes, in the etiology of Alzheimer's disease and in mechanisms underlying cancer. Recent studies have shown that over eighty-eight mitochondrial proteins are O-GlcNAcylated and that O-GlcNAcylation also not only regulates mitochondrial functions, but also that O-GlcNAc cycling is grossly abnormal in cardiac mitochondria from diabetic animals. These findings in mitochondria provide a molecular basis for mitochondrial dysfunctions in hyperglycemia associated with diabetes. Other recent studies have shown that O-GlcNAcylation plays a major role in neuronal functions, including synaptic functions and learning and memory. An inducible knockout of O-GlcNAc transferase in the excitable neurons of adult mice (driven by the □CAMKII promoter), surprisingly leads to a morbidly obese mouse within only about three weeks. These mice have a pronounced satiety defect mostly caused by defects in thyroid releasing hormone (TRH) producing neurons of the PVN of the hippocampus, a region of the brain long known to control appetite. Thus, O-GlcNAc is not only a nutrient sensor in all cells, but serves to regulate appetite control by the brain. Supported by NIH R01DK61671; P01HL107153; and N01-HV-00240. *Dr. Hart receives a share of royalty received by the university on sales of the CTD 110.6 antibody, which are managed by JHU.*

Development and Application of MS-based Strategies for Qualitative and Quantitative Analysis of Protein and Peptide PTMs

Lingjun Li, School of Pharmacy and Department of Chemistry, Neuroscience Training Program, University of Wisconsin-Madison, 777 Highland Avenue, Madison, Wisconsin 53705, USA

Protein post-translational modifications (PTMs) often play key roles in many biological processes. Alterations of these PTMs could significantly affect the function of target proteins or peptides, and are known to be associated with many diseases. Mass spectrometry (MS) has become a central tool to characterize these PTMs due to its chemical specificity, ability to analyze complex mixtures and the speed. This presentation will focus on our recent progress in the development and application of MS-based tools for large-scale glycoproteomic analysis in Alzheimer's disease (AD) and phosphoproteomic analysis in neuropathic pain. The remaining challenges associated with the study of PTMs and on-going efforts on the use of isobaric tagging strategies for quantitative PTM analysis will be discussed as well.

In an effort toward biomarker discovery in AD patients, we conducted a large-scale comparative glycoproteomic analysis via lectin affinity chromatography to enrich glycoproteins from cerebrospinal fluid (CSF) samples collected from control, mild cognitive impairment (MCI) and AD groups. Preliminary results revealed 137, 145, and 132 glycoproteins in control, MCI and AD groups respectively. Among these proteins, 75 identifications showed increasing or decreasing trend from control-MCI-AD, which could be potential biomarker candidates. To further characterize glycan structures in a quantitative manner, we employ a capillary electrophoresis (CE)-ESI MS platform for multiplexed quantitation enabled by carbonyl-reactive aminoxy tandem mass tag (aminoxyTMT) reagents.

In addition, we performed the first large-scale phosphoprotein analysis of a neuropathic pain model, elucidating the dynamic phosphoproteomes of the dorsal root ganglion (DRG) and spinal cord (SC) tissues under nerve injury, which are essential to the establishment of chronic pain. Initial results identified and quantified nearly 10000 distinct phosphoprotein isoforms. Several highly upregulated and downregulated proteins were kinases and phosphatases, suggesting important role of phosphorylation in neuropathic pain. The pulsed radiofrequency (PRF) treated condition revealed down-regulation of phosphorylation in several classes of trans-membrane ion transporters.

Towards Top-Down Proteomics

Ljiljana Paša-Tolić, Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA

Combinatorial post-translational modifications (PTMs), signal peptide cleavages, proteolytic processing and site mutations are all important biological processes that largely go undetected in traditional bottom-up proteomic analyses. While several PTMs are successfully identified using bottom-up methods, information including stoichiometry of modifications on a single protein, or presence of a combination of multiple modifications on a single proteoform[1] is practically impossible to infer from peptide-level data. Hence, the potential information gleaned from top-down (i.e. intact protein) studies, or through integration of top-down and bottom-up approaches,[2] is vast and is rapidly becoming an important avenue for proteomic studies.

Recent advances in MS instrumentation, separation, and bioinformatics significantly increased the throughput of top-down proteomics, allowing the identification of hundreds of intact proteins and their isoforms.[3] However, most of these efforts involve additional sample pre-fractionation steps, which are often labor intensive, require large sample sizes, and are inadequate in terms of quantitation. To tackle these challenges, we have optimized commercially available LCMS platforms for high-throughput, comprehensive and sensitive top-down quantitative analysis. This approach has been successfully applied for broad characterization of modifications on intact Salmonella proteins potentially relevant to pathogen biology;[4] characterization of the native forms of human salivary proteins potentially relevant to oral salivary diagnostics;[5] and insights into the underexplored mechanism of epigenetic control of gene expression for e.g. generating profitable bioactive compounds in fungus. Examples featured here highlight the complexity of comparing peptide abundance values in the context of protein abundance, and suggest that future top-down studies may be required for comprehensive analysis of biological processes.

1. Smith LM et al. *Nat Methods*. 2013 Mar;10(3):186-7.
2. López-Ferrer D et al. *Mol Cell Proteomics*. 2011, 10(2):M110.001479.
3. Tipton JD et al. *Anal Chem.* 2012 84:2111-7; Tran JC et al. *Nature*. 2011 480(7376):254-8; Catherman AD et al., *Biochem Biophys Res Commun.* 2014, 445(4):683-93.
4. Ansong C et al. *Proc Natl Acad Sci U S A*. 2013, 110(25):10153-8.
5. Wu S et al., *Proteomics*. 2014, 14(10):1211-22.

A New Frontier in Proteomics: Identifying Proteoforms and Elucidating Proteoform Families from Measurements of Intact Mass and Lysine Count

Lloyd M. Smith*, Michael R. Shortreed, Mark Scalf, Rachel A. Knoener, Anthony J. Cesnik, and Brian L. Frey
Department of Chemistry and Genome Center of Wisconsin, University of Wisconsin- Madison, Madison, WI 53706

The dominant paradigm of modern proteomics today is the "bottom-up" strategy, in which a mixture of proteins of interest is cleaved into peptides and analyzed by liquid chromatography/mass spectrometry (LC-MS). While the bottom-up strategy is powerful and widely practiced, the digestion of the proteins into peptides means that information as to the protein context within which that peptide is found is lost. Proteins produced from the same gene can vary substantially in their molecular structure: genetic variations, splice variants, RNA editing, and post-translation modifications (PTMs), all give rise to different forms of the proteins: these are referred to as "proteoforms". Knowledge of the proteoforms that are present in a system under study is absolutely essential to understanding that system, as the different proteoforms often have dramatically different functional behaviour, and regulation of their production is a central aspect of pathway control.

We are developing a new strategy for proteoform analysis, in which the determination of just two pieces of information for each proteoform, namely the accurate mass and the number of lysine residues contained, suffices to identify it. The accurate mass is determined by standard LC-MS analysis of the undigested protein mixture in an orbitrap mass spectrometer, and the lysine count is determined using a recently developed isotopic tagging method. A key enabling concept is a search strategy that reveals post-translationally modified protein variants. The strategy is demonstrated by elucidating hundreds of proteoform families present in yeast cell lysate. This simple and readily implemented new proteomic strategy provides an unprecedented view of the proteoforms present in biological systems, and will thereby make possible critical new insights into the functioning of biological systems and pathways.

Seeing Red and Going Green: Regulation of Eukaryotic Plasma Membrane Primary Transport Pumps by Phosphorylation, in Plants and Electric Fish

Michael R. Sussman, Department of Biochemistry and UW Biotechnology Center, UW-Madison

Quantitative mass spectrometic-based analysis of ser/thr/tyr phosphorylation is being utilized in my lab to investigate posttranslational modifications of plasma membrane proteins involved in regulating solute transport, hormone action and cell size. As the primary active transport system at the plasma membrane, all eukaryotes utilize a Mr=100,000 dalton polypeptide that encodes a polytopic hydrophobic membrane protein with ten transmembrane domains, that hydrolyzes ATP and pumps either a sodium (animals) or proton (plants, fungi) into the extracellular space. For plant studies, we use the biochemical and genetic technologies available with the model plant, *Arabidopsis thaliana* (1). For animal studies, we use the electric organ of *Electrophorus electricus*, the strong voltage electric eel. This electric fish has been a mainstay for biochemical studies of plasma membrane transport proteins for many years, and to facilitate this work, we recently completed a draft of its genome sequence (2).

In both animals and plants, this enzyme's catalytic properties are under the control of protein kinases and phosphatases that modify ser/thr and tyr residues. In animals, this enzyme pumps three sodium ions out, and two potassium ions in. In plants, it only pumps one proton out. This difference in stoichiometry creates large differences in transport capabilities. For example, the plant proton pump has a much larger reversal potential and can generate membrane potentials higher than minus 250 millivolts (inside negative) while the animal enzyme

does not generate transmembrane potentials above 150 millivolts (inside negative). In both kingdoms, this primary active transporter, regardless of its cation, provides all of the energy that drives solute transport. They also participate in poorly defined aspects of signaling pathways that regulate cell size and and in response to changes in the extracellular ionic milieu. In the electric fish, the sodium pump is asymmetrically located in specialized cells derived from muscle, the electrocytes, that are arranged on top of each other in large stacks, much like batteries in a flashlight, in order to multiply a transcellular voltage, for powering their electric organ. The 'firing' of each electrocyte is controlled by a nerve and altogether, the organ in the strong voltage electric eel can create an electric discharge with enough power to kill any other animal it meets. In higher plants, while there is no similar electric organ, each cell can generate a very large membrane potential and this is used to drive cell expansion. Changes in cell elongation are a key element of plant growth, that allows them to find sunlight and escape predation, even though they cannot move about within the soil. In this talk I will demonstrate how phosphorylation regulates these two enzymes from different phylogenetic kingdoms, and contrast nature's exquisite ability to use billions of years and lots of cell divisions, to create unusual adaptions for survival.

1. Haruta M, Sabat, G., Stecker, K., Minkoff, B.B. and Sussman, M. R. 2014 A peptide hormone and its receptor protein kinase regulate cell expansion in *Arabidopsis thaliana*. Science 343:408-411.
2. Gallant JR, Traeger LL, Volkening JD, Moffett H, Chen PH, Novina CD, Phillips GN Jr, Anand R, Wells GB, Pinch M, Güth R, Unguez GA, Albert JS, Zakon HH, Samanta MP, Sussman MR. 2014 Genomic basis for the convergent evolution of electric organs. Science. 344:1522-5.