3.1 Quantitative Analysis of Proteome Localisation and Dynamics

A. Lamond

Wellcome Trust Centre for Gene Regulation and Expression, MSI/WTB Complex, University of Dundee, Dundee, Scotland, United Kingdom

We are studying the functional organization of mammalian cell nuclei using a dual strategy that combines mass spectrometry (MS) based proteomics with live cell fluorescence imaging (see www.LamondLab.com). This applies two distinct but complementary quantitative techniques to analyse the same biological problem, providing a rigorous approach where potential artifacts or limitations of one method are avoided in the complementary approach and vice versa. The quantitative proteomic methods involve metabolic labeling of cellular proteins in cultured cell lines with the amino acids lysine and arginine containing heavy isotopes such as $^{13}$C and $^{15}$N. The quantitative imaging experiments, including time-lapse microscopy, FRAP, FLIP, FLIM and FLIM-FRET, are performed on mammalian cell lines stably expressing one or more fluorescent protein-tagged reporters. Both the proteomics and microscopy methods are used to study the same stable cell lines, allowing a direct comparison of the resulting data from both techniques.

We have used this dual strategy to characterize in detail the molecular composition of nucleoli under different metabolic and growth conditions and at specific stages of cell cycle progression (see http://lamondlab.com/noppdb/). We have developed a MS-based proteomics strategy to perform quantitative analyses of subcellular protein localization – “spatial proteomics” – including the analysis of protein turnover rates in separate cell compartments. This provides a new approach for annotating the spatial organization of the proteome and for measuring how this changes in response to inhibitors and different cell growth conditions. We have also developed quantitative MS-based approaches for identifying specific protein-protein interactions. These strategies provide a general approach for characterizing the composition, dynamic properties and interactions of either cell organelles or multi-protein complexes.

3.2 Post-Translational Adenosine Monophosphate (AMP) Modification of Proteins

C. A. Worby1,2,3,9, S. Mattoo4,9, R. P. Kruger5, L. B. Corbeil6,7, A. Koller6, Juan C. Mendez6, B. Zekarias6, C. Lazar1,2,3, and Jack E. Dixon1,2,3,4

Departments of 1Pharmacology, 2Cellular and Molecular Medicine, 3Chemistry and Biochemistry, and 4Howard Hughes Medical Institute, University of California, San Diego, La Jolla, CA; 5Department of Biological Chemistry, University of California, Ann Arbor, MI; 6Department of Pathology, University of California, San Diego Medical Center, San Diego, CA; 7Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA; 8Department of Pathology, Stony Brook University, Stony Brook, NY

Eukaryotic cells have devised different strategies to regulate signaling pathways. The best known modification is phosphorylation, which attaches a phosphate group to serine, threonine or tyrosine residues in proteins, thereby regulating their activities. Here, we describe a new modification: the addition of adenosine mononucleoside phosphate (AMP) on tyrosine residues. AMP addition to Rho GTPases by the Fic domain containing secreted surface antigen IbpA of the respiratory pathogen Histophilus somni leads to cytoskeletal collapse in host cells (1). Specifically, incubation of purified Rho GTPases (RhoA, Rac1 and Cdc42) with GST-tagged and purified Fic domain of IbpA in the presence of $\gamma^{32}$P-ATP, but not $\gamma^{32}$P-ATP, allows transfer of the $\gamma^{32}$P-label to RhoA, Rac1 or Cdc42, thus indicating the addition of AMP versus a phosphorylation event. Unlike VopS, another Fic domain containing protein from V. parahemolyticus, that modifies AMP on threonine residues (2), mass spectrometric analysis of IbpA-treated Rho GTPases show that the IbpA Fic domain adds an AMP to a conserved tyrosine residue in the switch I region of Rho GTPases. In addition, we show that the only human protein containing a Fic domain, HYPE (Huntingtin Yeast-interacting Protein E), also has the ability to add AMP to tyrosine residues in Rho GTPases in vitro. Thus, we identify Fic domain containing proteins as a new class of enzymes that mediate not just bacterial pathogenesis, but also a previously unrecognized eukaryotic post-translational modification that may regulate key signaling events.

Interestingly, threonine and tyrosine AMP modified peptides behave similarly in the mass spectrometer as threonine and tyrosine phosphorylated peptides: whereas threonine modified peptides undergo neutral loss of AMP (plus 18Da) on fragments upon activation of the peptide, peptides modified with AMP on tyrosine mostly stay intact, partially losing adenine as well as adenosine. In addition, AMP-Tyr modified peptides are only identifiable in an ion-trap CID fragmentation experiment, as fragmentation in an HCD cell (or CID in a QSTAR) will lead to a strong signal for adenine and only weak fragmentation peaks of the peptide backbone.

References

1. Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., Zekarias, B., Lazar, C., and Dixon, J. E. (2009) The fic domain: regulation of cell signaling by adenylylation. Mol. Cell 34(1), 93–103.
2. Yarbrough, M. L., Li, Y., Kinch, L. N., Grishin, N. V., Ball, H. L., and Orth, K. (2009) AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. Science 323(5911), 269–272.
3.3 Dissecting the Structure of the Human Spliceosome by Looking at Its Pieces

P. Coltri\textsuperscript{1}, J. Ilagan\textsuperscript{1}, R. J. Chalkley\textsuperscript{2}, A. L. Burlingame\textsuperscript{3}, and M. S. Jurica\textsuperscript{1}

\textsuperscript{1}Department of Molecular, Cell and Developmental Biology and Center for Molecular Biology of RNA, University of California, Santa Cruz, CA; \textsuperscript{2}Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA

Pre-mRNA splicing is the removal of the non-coding introns that interrupt most gene transcripts and serves an essential step in eukaryotic gene expression. The cellular machinery responsible for splicing, termed the spliceosome, is a large protein/RNA macromolecular complex comprised of five structural RNAs and over 100 individual polypeptides. The human complex assembles and functions via a progression of structural intermediates that are not yet fully characterized. The dynamics and complexity of the spliceosome are long posed challenges to detailed biochemical and structural studies that will provide insight into the spliceosome’s molecular mechanisms. In particular, isolating distinct conformations of this moving target in the amounts needed for standard biochemical and structural analyses is not simple. We made a key advance in this regard with our development of a substrate-based affinity method to isolate human spliceosomes arrested midway through splicing catalysis (C-complex). Initial mass spectrometry analysis of this complex identified over 200 proteins, \textasciitilde100 of which were specific to splicing. Using cryo-electron microscopy (cryo-EM) and single particle reconstruction techniques, we solved the structure of C-complex spliceosomes to 30 Å resolution. This model represents an important first step in visualizing the structure of the spliceosome. However, before we can more fully interpret the model in functional terms we must answer questions regarding which components of the spliceosome are visualized/represented in our model and where they are located in the structure. Currently, we are finding ways to take the spliceosome apart and then examining the resultant pieces. Mass spectrometry analysis is critical for defining the protein composition of the pieces, enabling us to define interactions that underpin the spliceosomes’ architecture. We have examined the contribution of exon sequences in the composition and structure of C-complex and are now looking at the proteins that tightly associate with the intron vs. the region of the upstream exon poised for ligation. In addition to these studies, we have made progress in using chemical modification in conjunction with mass spectrometry to identify regions of proteins that are located at the surface of the spliceosome. This work will allow us to begin localizing these proteins’ positions in the complex. By combining the results of these studies with our structural investigation of the spliceosome, we are on the path to assembling a more detailed model of this critical cellular machine.

3.4 Protein Complexes and Functional Pathways in S. cerevisiae and E. coli

M. Babu\textsuperscript{1}, G. Butland\textsuperscript{2}, J. J. Díaz-Mejía\textsuperscript{1,4}, P. Hu\textsuperscript{1}, S. Pu\textsuperscript{2}, G. Moreno-Hagelsieb\textsuperscript{4}, S. C. Janga\textsuperscript{1}, S. Wodak\textsuperscript{2,6}, A. Emili\textsuperscript{1,2}, and J. Greenblatt\textsuperscript{1,2}

\textsuperscript{1}Banting and Best Department of Medical Research, \textsuperscript{2}Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada, \textsuperscript{3}Life Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, \textsuperscript{4}Department of Biology, Wilfrid Laurier University, Waterloo, ON, Canada, \textsuperscript{5}Hospital for Sick Children, Toronto, ON, Canada

We have used TAP-tagging and affinity-purification to sort the soluble proteins of S. cerevisiae into complexes. We combined this with systematic synthetic genetic interaction analysis for non-essential gene deletion mutants and essential gene hypomorphs, using the synthetic genetic array (SGA) approach, for genes related to nuclear processes. More recently, we have extended the yeast protein interaction network by focusing on the predicted yeast membrane proteins, purifying each protein three times in the presence of different detergents. We are testing the co-functionality of proteins in various membrane-associated protein complexes by comparing our protein complex data with synthetic genetic interaction data and assessments of the effects of the various proteins in a complex on the morphology of the intracellular compartment in which that complex is located.

We have also used dual affinity tagging followed by affinity purification and mass spectrometry to sort the soluble proteins of E. coli into protein complexes. Although our initial focus was on essential, evolutionarily conserved proteins, we have focused more recently on proteins of unknown function (functional orphans). We integrated our protein-protein interaction network with systematic genome context inferences to derive a probabilistic network of functional inferences encompassing almost all E. coli proteins (98%) and to assign about 57% of the orphans to discrete functional neighborhoods with high confidence. Many of these functional inferences were then confirmed by genome-scale phenotypic assessments. Functional pathways can be derived by systematically identifying genetic interactions, or epistasis, which tends to occur between genes involved in parallel pathways or interlinked biological processes. We have therefore developed a quantitative screening procedure, eSGA (E. coli synthetic genetic arrays), for monitoring bacterial genetic interactions based on conjugation of E. coli deletion or hypomorphic strains to create double mutants on a genome-wide scale. The patterns of synthetic lethal- or sickness (aggravating genetic interactions) we observe for certain double mutant combinations provide information about functional relationships and redundancy between pathways and enable us to group E. coli genes into functional modules.
3.5

N-Terminomics: High Confidence, Broad Dynamic Range Coverage Utilizing Novel Polymers for Proteomics Reveals the Functional State of the Proteome

C. M. Overall

UBC Centre for Blood Research, University of British Columbia, Vancouver, British Columbia, Canada

The nature of a protein’s N-terminus, its modifications and sequence has profound impacts on the function and localisation of proteins. Moreover, all proteomes are moulded by proteolysis and in all cases this changes the function of a protein, for example in enzyme and protein activation, inactivation, conversion to antagonists, as triggers for secretion, cell surface shedding and finally clearance. Therefore, to functionally annotate proteins, the N and C terminal peptides of proteins must be determined. To focus on these peptides, dedicated techniques are required. Hence, these semitryptic peptides are all too often overlooked in proteomics analyses.

We have developed novel polymers that target primary amine groups on peptides that are invaluable for such proteomics analyses because of their high derivatisation, excellent solubility, no non-specific binding properties, low cost and easy synthesis. Using these polymers in a new procedure termed TAILS (Terminal Amine Isotope Labelling of Substrates) we report a new proteomic and bioinformatics pipeline to rapidly identify natural N-termini and protease cleaved neo-termini of protein substrates after polymer enrichment. MS/MS both identifies the N-terminal peptide and both the substrate and sequence of protease cleavage sites in the same experiment. For most proteins multiple peptides are so identified enabling robust protein identification through multiple peptides. For proteins with single peptides identified, a new statistical analysis enables high confidence protein identification. The key to identifying specific protease substrates is the use of isotopic labelling of all primary amines in order to subtract background proteolysis that is always present. This can be achieved by dimethylation and iTRAQ labelling on primary amines in 8-plex analyses.

We applied TAILS for quantitative N-terminome analysis and for the global analysis of proteolysis in skin inflammation induced by TPA (12-O-tetradecanoyl-phorbol-13-acetate). First, we developed and successfully tested a mass spectrometry-compatible protein isolation and purification method for total skin lysates. Next, we combined this method with TAILS analysis to determine both the skin proteome and skin N-terminome and their perturbations in inflammation. Including wild-type and matrix metalloproteinase (MMP) 2 knockout mice in this multiplex approach allowed us to further identify novel bioactive substrates of this important family of inflammatory matrix metalloproteinase. Thereby, we identified 1,972 proteins with high confidence from murine skin samples with 84 being significantly up-regulated in TPA treated skin including known inflammatory markers such as acute phase proteins and components of the complement system. By TAILS we identified 1,677 N-terminal peptides for 1,032 proteins including 621 that had also been detected prior to N-terminal enrichment. Importantly, among the 411 proteins only identified after enrichment for N-termini were low abundance chemokines like the small inducible cytokines B5 (LIX) and macrophage inflammatory protein 2 (MIP2). As expected, the N-termini of these proteins were also included in a subset of 312 N-terminal peptides assigned to 184 proteins significantly induced by TPA treatment with a statistically significant enrichment of inflammation-related categories by Gene Ontology (GO) analysis. Notably, the analyses were neither skewed by proteins that are highly abundant in skin, such as keratin and filaggrin, nor by serum proteins (only 23 identified).

Hence, N-terminomics analyses using negative peptide selection enables broad proteome coverage with high dynamic range of complex proteomes.