Mass Spectrometry–Based Proteomics and Network Biology

Ariel Bensimon,[1,2] Albert J.R. Heck,[1,3] and Ruedi Aebersold[1,2,4]

[1]Department of Biology, Institute of Molecular Systems Biology, and [2]Competence Center for Systems Physiology and Metabolic Diseases, ETH Zurich, CH 8093, Switzerland; email: bensimon@imsb.biol.ethz.ch, aebersold@imsb.biol.ethz.ch

[3]Biomolecular Mass Spectrometry and Proteomics, University of Utrecht, and Netherlands Proteomics Center, 3584 CH Utrecht, The Netherlands; email: a.j.r.heck@uu.nl

[4]Faculty of Science, University of Zurich, CH 8006 Switzerland

Annu. Rev. Biochem. 2012. 81:379–405

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

This article's doi: 10.1146/annurev-biochem-072909-100424

0066-4154/12/0707-0379$20.00

Keywords

systems biology, quantitative proteomics, protein interaction network, protein-signaling network, posttranslational modification, data-driven modeling

Abstract

In the life sciences, a new paradigm is emerging that places networks of interacting molecules between genotype and phenotype. These networks are dynamically modulated by a multitude of factors, and the properties emerging from the network as a whole determine observable phenotypes. This paradigm is usually referred to as systems biology, network biology, or integrative biology. Mass spectrometry (MS)–based proteomics is a central life science technology that has realized great progress toward the identification, quantification, and characterization of the proteins that constitute a proteome. Here, we review how MS-based proteomics has been applied to network biology to identify the nodes and edges of biological networks, to detect and quantify perturbation-induced network changes, and to correlate dynamic network rewiring with the cellular phenotype. We discuss future directions for MS-based proteomics within the network biology paradigm.

Contents

MOLECULES TO NETWORKS

Much of life science research has been focused on understanding the complex relationship between genotype and phenotype. Specifically, research has addressed the fundamental questions of how, when, and where the information encoded in the genome of an organism is expressed and modulated by external (e.g., environmental) or internal (e.g., genomic) factors to generate a specific phenotype.

Over the past decades, such studies have been carried out within the "one gene-one protein-one function" paradigm referred to as the "molecular biology paradigm," which arose from the classical work of Beadle & Tatum (1) on amino acid metabolism in *Neurospora*. This paradigm makes two important assumptions that have dominated the thinking of generations of experimental biologists and guided the development of the techniques of molecular biology. First, it postulates a direct link between gene and protein function, implying that knowledge of all the genes and their translation products can explain biological function. Second, it orders individual proteins and their associated functions in linear pathways, implying that every function "downstream" is affected by an upstream block, whereas every function "upstream" is unaffected by a downstream block (**Figure 1*a***, *left*). In the genomic age, powerful technologies have been developed to support research at a global scale within the molecular biology paradigm. These include genome sequencing to identify all protein-coding genes of a genome (2, 3); proteomic methods to identify and quantify the proteins in a biological sample (4); genomic engineering (5); and gene knockout, RNAi technologies, and small-molecule inhibitor screens to inhibit or manipulate specific functions and to identify upstream and downstream events (6). Genome-wide RNAi screens that essentially search the whole genome space have been particularly popular. They are often applied to link genes to phenotypic readouts on a global level (for example, References 7–9). Overall, the technologies to identify, quantify, mutate, and interfere with the expression levels of any conceivable gene or protein of a species have reached a very high level of maturity.

In spite of these impressive technical advances and their wide and successful application, it has generally remained challenging to establish genotype-phenotype links. For example, with the exception of relatively few single gene defects with high penetrance, the molecular basis of most disease phenotypes turned out to be more complex and remain to be determined (10, 11). The reasons for these difficulties are likely conceptual rather than merely technical. The molecule-centric, single

RNAi: small RNA molecules that interfere with messenger RNA and thus prevent the translation of a protein

directional pathway-based paradigm, focusing on the properties of molecules, has turned out to be limited because it has neglected contextual relationships, such as cross talk between linear pathways that were considered to operate in isolation of one another.

Recently, a new paradigm has been emerging, typically referred to as systems biology, network biology, or integrated biology, that takes into account contextual relationships (12–15). In this network view, each node represents a molecule of interest, such as a gene, any of its products, or smaller molecules, such as cofactors, messenger molecules, and metabolites. The edge between two nodes represents a relationship, such as a physical interaction, an enzymatic reaction, or a functional connection. Although the molecule-centric paradigm is concerned with the network nodes, the new paradigm is concerned with the network nodes and edges, placing networks of interacting molecules between genotype and phenotype (**Figure 1*a***, *right*). It assumes that the structure and topology of such networks are an expression of the genomic information, that the networks are dynamically modulated at different timescales by external (e.g., environmental factors) or internal (e.g., genomic alterations) perturbations, and that the properties of the entire network determine the phenotype.

The network biology paradigm has several important implications. First, it requires a different, more integrative view of biological processes, as the contextual relationships between molecules move to the forefront. Second, network biology provides new opportunities for and critically depends on new experimental and computational approaches, including methods to visualize networks, methods to infer network topology and structure, and methods to simulate and model the dynamic behavior of networks and their phenotypic consequences. Thus, network biology has been stirring novel technologies that focus on the measurement of contextual relationships of molecules, rather than on simply enumerating molecules in a catalog format. In this review, we discuss the current state of mass spectrometry (MS)–based proteomic approaches that support network biology.

MASS SPECTROMETRY–BASED PROTEOMICS

The main goal of proteomics is the detailed characterization of the proteome. In the molecular biology paradigm, the focus has been on the comprehensive identification and characterization of protein sequences, including their posttranslational modifications (PTMs), and on the comprehensive quantification of the protein components of a biological sample.

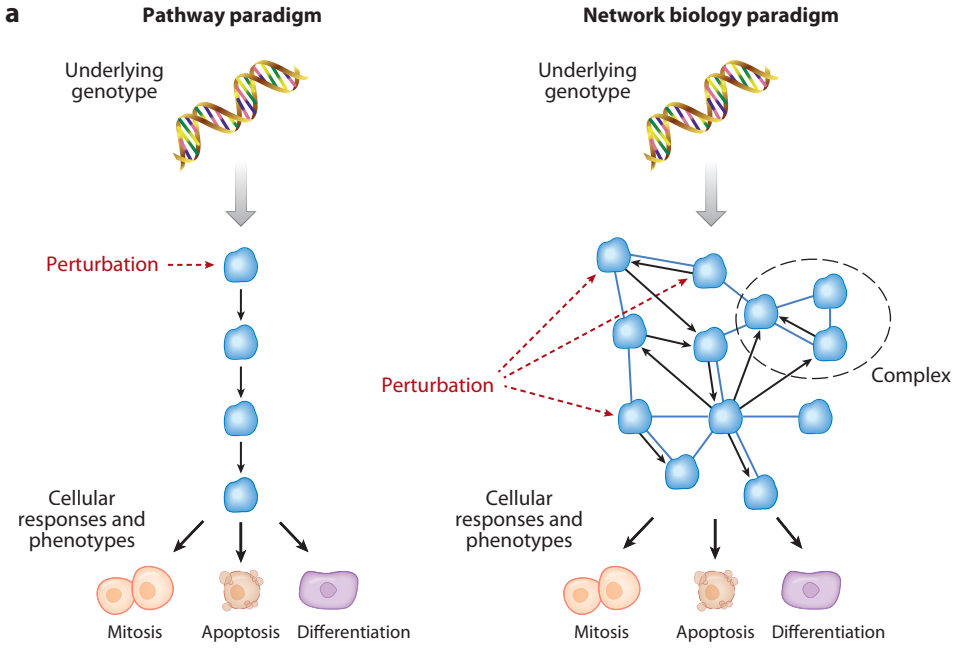
Most proteomic studies rely on tandem mass spectrometry as the core technology, specifically on a method referred to as bottom-up proteomics. In bottom-up proteomics, protein samples extracted from cells or tissues are digested into peptides. Peptides in the sample are then separated, typically by liquid chromatography, ionized, and transferred into the mass spectrometer, where peptide fragment ion spectra are recorded. Fragment ion spectra are the currency of information in bottom-up proteomics, as they can be assigned to peptide sequences from which the corresponding proteins are inferred. Fragment ion spectra are also used to detect modified amino acid residues and to identify and locate modifications within the peptide sequence. Peptide ion signals can also be used to infer the quantity of a sample peptide or protein (16, 17). For every step of the process, including sample preparation and fractionation, MS data acquisition, quantification, and data analysis, multiple methods and tools have been developed and reviewed extensively (4, 16–19). This also applies to the MS instrumentation, which enjoys a continued increase in performance in regard to mass accuracy, sensitivity, and analytical robustness (16, 17).

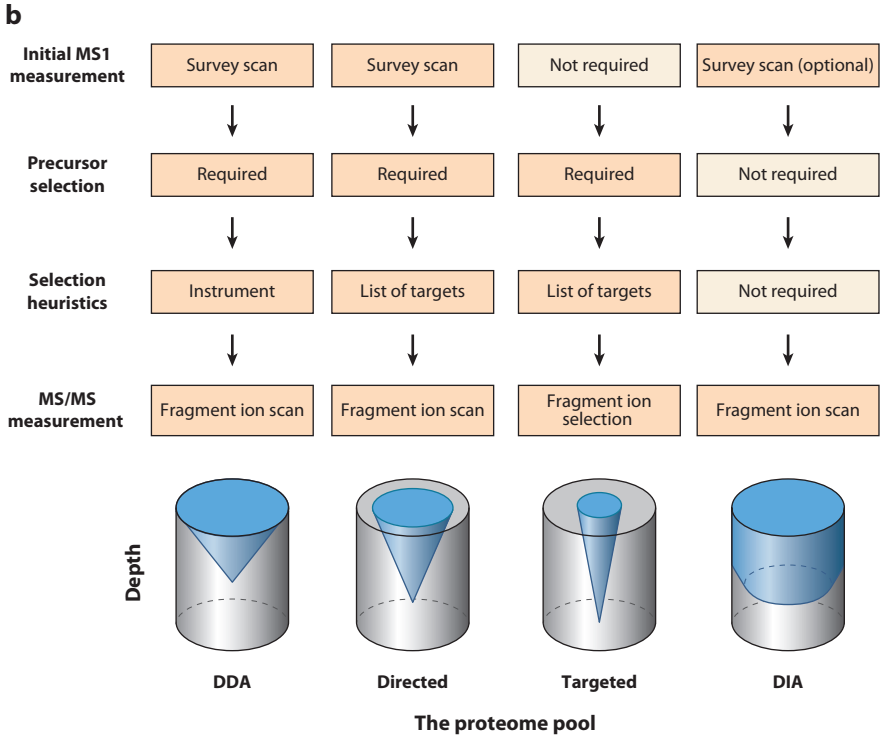
From an extensive menu of available options for each procedure step, individual choices have been combined into different workflows and MS strategies, each addressing different types of biological inquiries (16, 17, 19). These can be applied to various samples, such as whole proteomes, or enriched fractions, such as

MS: mass spectrometry

Posttranslational modifications (PTMs): the enzymatic covalent addition of a molecular entity to a protein, or removal of a molecular entity, or the irreversible change in protein sequence

Liquid chromatography: physical separation of analytes by the differential partition of each analyte between a mobile and a stationary phase in a column





phosphoproteomes. The most frequently used strategy is referred to as shotgun or discovery proteomics. There, precursor ions are detected in a survey scan and selected automatically using a simple heuristic via a process referred to as data-dependent analysis. This strategy results in datasets that can identify vast numbers of proteins and enables quantitative comparison between samples, either with stable isotope labeling or without labeling, in an approach known as label-free quantification (4, 20–22). Shotgun proteomics does not require any prior knowledge of the composition of the sample, and thus each protein in every sample analyzed is newly discovered. In directed proteomics, precursors are only selected for fragmentation if they are detectable in a survey scan and present on a list of predetermined precursor ions, i.e., an "inclusion list" (23, 24). This strategy results in datasets that identify and quantify specific, predetermined segments of a proteome at a higher level of reproducibility, compared to discovery proteomics. In targeted proteomics, only predetermined peptides are selected for detection and quantification in a sample. The main mass spectrometric method supporting targeted proteomics is selected reaction monitoring (SRM) also referred to as multiple reaction monitoring (25, 26). In SRM, specific mass spectrometric assays are generated a priori for each targeted peptide, and these assays are then used to selectively detect and quantify analytes in multiple biological samples (27). This method can generate highly reproducible and accurate datasets of a small, preselected fraction of a proteome (typically one to a few hundred peptides) at a wide dynamic range (17, 28). Finally, with recent advances in instrumentation, a fourth strategy referred to as data-independent analysis (29–32) is emerging in which no selection of precursor ions occurs, i.e., the fragmentation of all precursors is attempted for each sample, the analysis of which can benefit from the availability of extensive spectral libraries (27, 33). Each of these strategies captures a different subset of the "total proteome space" (**Figure 1*b***), balancing trade-offs in comprehensiveness, reproducibility and selectivity, sensitivity, accuracy, and dynamic range (17, 34).

Proteomic studies in the molecular biology paradigm have, for the most part, strived to increase coverage of the discovered, characterized, and quantified proteome (34). This has been technically challenging to accomplish, but it is conceptually simple and largely achievable using the discovery proteomics strategy.

Stable isotope labeling: a technique in which samples are metabolically or enzymatically labeled with stable isotopes of different masses

Figure 1

(*a*) Within the one gene-one protein-one function paradigm, referred to as the molecular biology paradigm, individual proteins and their associated functions are ordered in linear pathways, implying that every function downstream is affected by an upstream block (*left*). In the network biology paradigm, networks of interacting molecules are placed between genotype and phenotype. Each node (represented in *blue*) represents a molecule of interest, such as a gene, any of its products, or smaller molecules, such as metabolites. The edge between two nodes represents any type of relationship, such as protein-protein interactions (*blue lines*), enzyme-substrate relationships (*black arrows*). Both nodes and edges may be subject to perturbation (such as stimulus, inhibition, knockout, etc.). (*b*) Several prominent mass spectrometry (MS)–based strategies are available. The most frequently used strategy is referred to as shotgun or discovery proteomics. This strategy results in datasets that can identify vast numbers of proteins contained in biological samples but is more likely to reproducibly detect the most abundant proteins. In directed proteomics, specific, predetermined segments of a proteome are identified and quantified at a higher level of reproducibility, compared to discovery proteomics. Targeted proteomics generates highly reproducible and accurate datasets of a small, preselected fraction of a proteome (typically one to a few hundred peptides) with high detection sensitivity and dynamic range. With recent advances in instrumentation, a fourth strategy referred to as data-independent analysis (DIA) is emerging in which the identification of all proteins is attempted for each sample. Abbreviation: DDA, data-dependent analysis.

PPI: protein-protein interaction

Yeast two-hybrid (Y2H) screening: a genetic technique used to screen for physical interactions between pairs of proteins

Notable projects have achieved a high degree of coverage for proteomes (35–40) and selected PTMs (41–43), have measured quantitative changes in protein and PTM abundances (44–48), and have estimated the absolute cellular concentrations of significant fractions of proteomes (36, 40, 49, 50). In many of these studies, the main result is a list of abundance-modulated proteins or modified peptides. Typically, one or a few of the list's elements are followed up by classical biochemical or cell biology methods, a strategy that may be successful but ultimately not satisfactory because most of the information collected in the proteomic screen remains unused. Alternatively, the generated lists are subjected to analysis by gene ontology enrichment tools, pathway and signal transduction databases, or protein-protein interaction (PPI) databases (51, 52). The aim of these analyses is (*a*) to relate the contents of the list to prior knowledge in the form of protein functional classes or pathways or (*b*) to provide a visual concept of cellular processes, e.g., in the form of a network (53). We see these analyses as post-measurement network approaches in a molecular biology paradigm, rather than studies motivated a priori by network biology. As such, although they may provide valuable knowledge, this review does not further focus on such studies.

APPLYING MASS SPECTROMETRY–BASED PROTEOMICS TO NETWORK BIOLOGY

In the network biology paradigm, the technologies available to identify and quantify molecules need to be applied to also determine or infer and quantify the edges of these networks, i.e., the wiring underlying cellular networks. Measuring network edges by large-scale "omics" studies has been addressed primarily by two approaches. The first direct approach uses the affinity between nodes to capture the interacting molecules and to thus directly measure an edge. This is exemplified by technologies such as chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) (54) or by affinity purification (AP)-MS (see below). Second, the indirect approach uses an assay to probe a relationship between two nodes and to thus infer an edge. This is exemplified by yeast two-hybrid (Y2H) screening or genetic interaction networks (55). The limited toolbox of experimental methods to detect and quantify network edges raises the important question of how dynamic networks can be best studied. In a seminal study, Ideker et al. (56) have addressed this question. They described a generic road map for network biology, consisting of the following steps, which can be applied iteratively: (*a*) defining an initial network of nodes and edges from prior information, (*b*) perturbing network components and integrating the experimental data obtained with the network model, and (*c*) refining the network model to better predict experimental data and phenotypes arising from the network. During the past decade, these steps have been explored by a multitude of experimental and computational strategies (12–15).

Although the network biology paradigm has progressed significantly at the conceptual level, it is still substantially bound by molecular biology data collection techniques. To support the road map outlined above, MS-based proteomics, like other data collection technologies, needs to be able to generate datasets that minimally fulfill the following criteria: (*a*) the data have to be complete, i.e., all the network nodes and edges should be measurable; (*b*) the data need to be reproducible, i.e., identical results should be obtained in each repeat measurement of a network; (*c*) the data have to be quantitative to detect dynamic changes of network components; and (*d*) the data need to be measurable at a reasonable throughput to allow iterations within a study. Clearly, in meeting some of these criteria, proteomics has lagged behind other genomic technologies.

In the following sections, we describe steps that have been taken in MS-based proteomics to advance our ability to investigate and compare networks via measurement of molecules.

We also briefly review several network-driven experimental and computational approaches that show promise for use with proteomic data. We structure our review to describe the ways MS-based proteomics has been applied to the generic road map of multidirectional network biology (**Figure 2*a***): The first tasks are to identify the network components (nodes and edges) as well as to describe network wiring and the principles of the network organization, the second task is to relate perturbations to quantitative measurements to describe dynamic network rewiring, and the third task is to correlate dynamic network wiring with the cellular phenotype. Finally, we discuss what lies ahead for MS-based proteomics in the path toward network models and their mechanistic analysis and refinement.

For clarity of presentation, we distinguish between two types of networks investigated by MS-based proteomics: protein interaction networks (PINs) and protein-signaling networks (PSNs). PINs are undirected networks, i.e., the edges show no preferred direction, whereas PSNs are directed networks, i.e., the edges have a preferred direction. Clearly, a cellular network is the sum of several such networks (**Figure 2*b***).

Importantly, the investigation of network wiring by MS-based proteomics can benefit from network data collected by other technologies. For example, genetic interaction networks examine the dependency in function between two genes, reported by a growth phenotype, and represent a unique case in which the edges are measured (as phenotypes) and nodes are not explicitly measured. Genetic interaction maps are often carried out under basal growth conditions (55, 57) but have also been used to capture context-specific interactions, such as a differential interaction map (58). Because of limited throughput at present, MS-based proteomics cannot be applied as a direct readout for genetic interaction studies, but it can use the complementary information such screens provide to highlight edges, which have functional significance within the context of describing network wiring.

DESCRIBING NETWORK WIRING

Here, we review how qualitative and quantitative MS-based proteomic techniques have been applied toward defining a qualitative network, i.e., to support statements such as "protein X interacts with protein Y," in the case of PINs, or "protein X is a target of protein Y," in the case of PSNs. Such a network is a prerequisite to provide a comprehensive and reproducible description of the underlying cellular network and to serve as an initial map for quantitative/dynamic analyses.

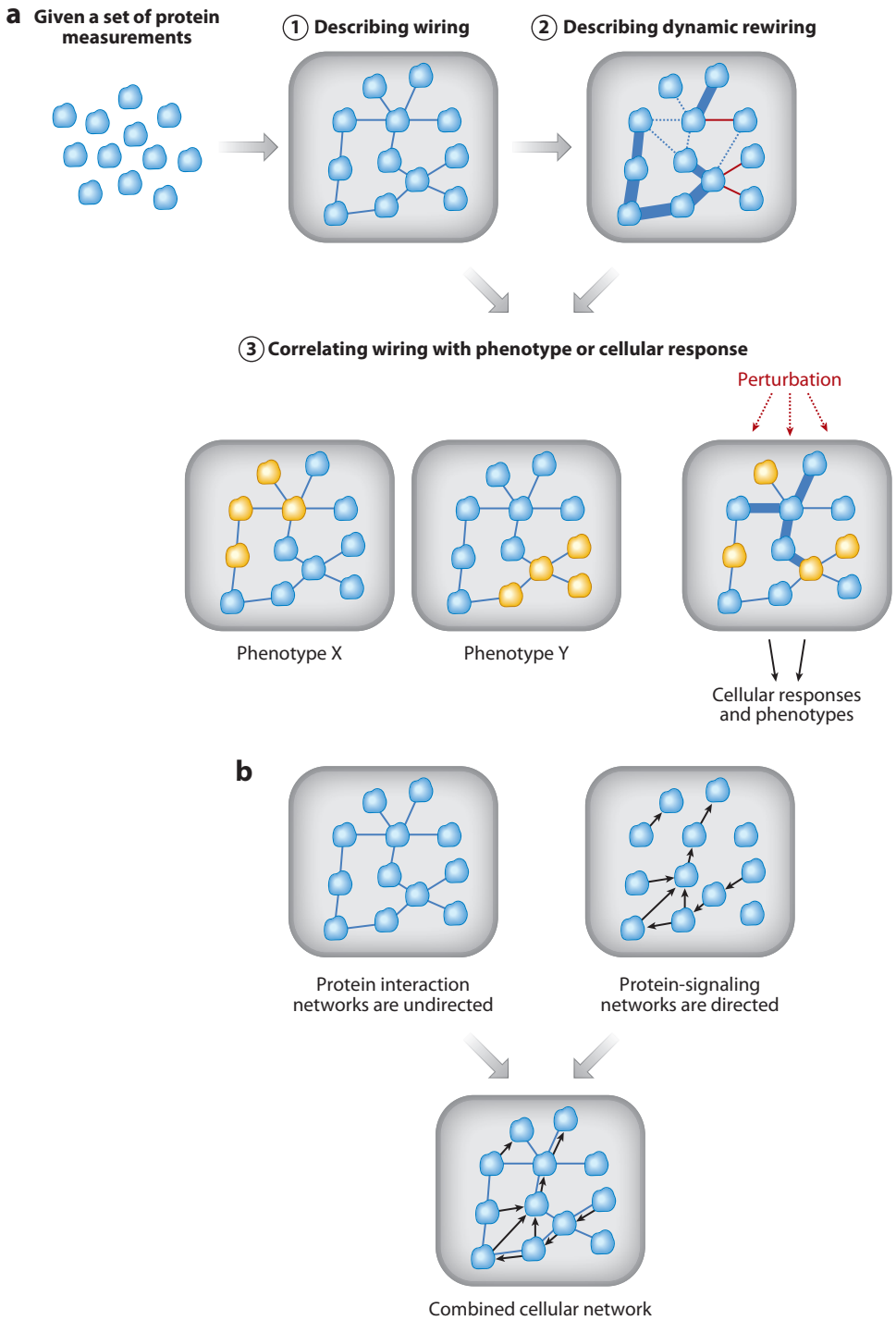
PIN: protein interaction network

PSN: protein-signaling network

Describing Protein Interaction Networks

PINs, exemplified by protein-protein interaction networks (PPINs), have been a major focus of interest in MS-based proteomics mainly because in these networks, in principle, both nodes and the edges linking them are directly measurable. The prototypical experimental approach to study such networks has been the use of a protein or other biomolecule as a "bait molecule" to capture and isolate "prey" proteins interacting with the bait, and to then identify these proteins (bait and preys) by MS. Although the following discussion mainly focuses on PPINs, proteins evidently also interact with other molecules that represent vital parts of the network (15). Studies focusing on interactions between proteins and other (cellular) molecules include those using RNA or drug molecules as bait to capture and identify interacting proteins (59–61). Conversely, proteins have been used as bait to capture copurified metabolites (62). Although these types of networks have not been investigated as routinely as PPINs, they provide essential information to understand the cellular network dynamics.

Prior to MS-based proteomics, the majority of the PPIN data were collected by Y2H screens, a proteomic technology with a high throughput. Though the quality of the Y2H PPI data has increased substantially over time (55), the method interrogates only binary interactions capturing only a subspace of the whole interactome (57). Therefore, AP of tagged



proteins of interest (POI) and identification of the copurified protein components by MS (AP MS) (63) have become a preferred method for the analysis of PPINs because the data it produces reflect more closely the actual multidirectional complexity of a PPIN in the cell. AP-MS data are then interpreted as a network of PPIs, whereby two or more copurified nodes are connected by edges. It should be noted that the majority of these experiments are performed in various cell lines and thus may not reflect correctly the cellular networks in tissues. Moreover, copurification does not necessarily indicate a direct physical interaction, and detailed information about the composition and structural topology of protein complexes is not directly apparent from AP-MS data. These issues can be addressed by cross-linking techniques and mass spectrometric analysis of intact complexes and have been reviewed elsewhere (63–66).

Inevitably as with every technology, AP MS comes with concerns regarding the quality of the obtained data. A first concern relates to the possibility of false-positive PPIs because in most species genetic manipulation is inefficient, and therefore, the tagged protein is often overexpressed from an expression vector and not expressed from the endogenous locus. This will remain an issue until techniques for tagging proteins by genetic manipulation of human cell lines (67) or tagging proteins in whole animals (68) is as feasible as tagging those for yeast cells. One method, suggested to circumvent this issue, is by expression of the protein from a bacterial artificial chromosome (69). Another way to avoid the problem is by immunoprecipitation of the endogenous protein whenever an antibody is available. For example, Malovannaya et al. (70) performed more than 3,000 immunoprecipitation experiments of endogenous proteins, including reciprocal immunoprecipitations, and uncovered in their vast dataset core complex modules, unique core "isoforms," and complex-complex interactions.

A second concern in this type of experiments is the identification of true and specific PPIs as opposed to nonspecifically copurified proteins. A simple approach to overcome this problem is the identification of common contaminants that copurify nonspecifically owing to their "stickiness" to the matrix or the tag affinity reagent (71, 72). Another way of addressing contaminants is by using quantitative interaction proteomics to compare an AP of a tagged POI to an AP that represents the unspecific background, for example, an untagged POI. A similar abundance of a protein in both purifications would indicate the protein is a contaminant (69, 71, 73, 74). On the basis of this concept, several scoring algorithms were developed to estimate the likelihood of an interaction by different metrics derived from a protein's spectral counts, thus extracting contaminants and identifying "true" interactors (75–77).

Thus far, much effort has been put into providing a description of the cellular PPIN. Several studies have addressed the landscape of PPINs through the analysis by AP MS of large groups of POIs: 75 human deubiquitinating enzymes (76); 32 human proteins linked to autophagy and vesicle trafficking (78); a

POI: protein of interest

AP MS: affinity purification coupled to mass spectrometry

Spectral count: counting the number of times a peptide is identified in a dataset as a proxy to its abundance

Figure 2

(*a*) Given a set of protein measurements by mass spectrometry (MS)-based proteomics, three tasks are faced with the generic road map of network biology: The first tasks are to identify the network edges and to describe network wiring, the second task is to relate perturbations to quantitative measurements to describe dynamic network rewiring (depicted as a variance in the weight of edges), and the third task is to correlate dynamic network rewiring with the cellular phenotype (depicted as *yellow nodes*). (*b*) Two types of networks are investigated by MS-based proteomics: Protein interaction networks are undirected networks, i.e., the edges show no preferred direction. Protein-signaling networks are directed networks, i.e., the edges have a preferred direction. The cellular network integrates a variety of those networks, which in proteomics are often investigated separately.

genome-scale analysis of protein complexes in the bacterium *Mycoplasma pneumonia* (79); and 276 kinases, phosphatases, regulatory subunits, and their scaffolds in yeast (80).

Frequently, PPIs are modulated by PTMs. Vermeulen et al. (81) identified histone mark "readers," proteins that interact with histone tail peptides trimethylated at various lysine residues. For selected readers, they inspected their interactomes by AP MS, the corresponding genomic binding sites by ChIP-seq, and their binding strength in the context of combinatorial histone modifications. This study certainly suggests that describing networks in the future will also have to account for the involvement of PTMs in forming these interactions.

Organizing networks into complexes. AP MS provides a description of the network wiring in the form of edges between nodes identified by MS, but it does not explicitly identify the composition of protein complexes. Therefore, after refining the PPIN for true interactions, the next challenge is to infer true protein complexes from AP-MS data. Experimentally, for any type of AP, a reciprocal AP of the identified copurifying proteins can greatly improve the confidence in the annotation of these relationships and refine the identification of putative complexes (70). Another approach, applied in yeast, is the AP MS of the same bait protein from various strains that were deleted for known interactors. This enabled capturing information about the association between every bait and prey, as well as every prey and the deleted known interactor (82), providing information about how the complex is assembled. In essence, both reciprocal AP and analysis by deletion tackle the identification of complexes by attempting to address the undirected edges in the PPIN from both ends (nodes). Computationally, PPI data have been used to infer complexes from the network topology either by various graph-based approaches (83) or by using MS quantitative data. This has been attempted by nested clustering (84), a biclustering approach, in which baits are first clustered by normalized spectral counts, and then preys are clustered within their respective bait clusters.

In sum, describing the wiring of PPI networks has seen significant advances in throughput and reproducibility of data collection. However, the technique is still far from being comprehensive, as with a few exceptions, such as yeast, for most species only small subsets of their respective proteomes have been thus far covered. Furthermore, PINs, other than PPINs, have been even more challenging to analyze.

Describing Protein-Signaling Networks

PSNs, exemplified by enzyme-substrate relationships, have been a second focus of interest in MS-based proteomics, particularly in the case of enzymes, which modify their substrate by a PTM. The addition or removal of a PTM to or from a peptide results in a change in mass detectable by MS. In PSNs, the interaction between pairs of an enzyme and a substrate tends to be very transient, and its significance lies in the directed transduction of a signal, represented as a directed edge. To describe the wiring of a PSN, we would require reproducible and comprehensive measurements of all proteins involved, thus determining all edges. In particular, the molecular context of these signaling events, in space and time, is also provided by a framework of other proteins, such as adaptor and anchor proteins (85). However, although, in a PPIN, an edge is measured by the concurrent measurement of the two connected nodes, in a PSN, this is not always feasible owing to the transient nature of interaction. Consequently, often only one node (a substrate, for example) of two connected nodes can be measured, and edges need to be inferred. Therefore, much more effort has been put in PSNs to first describe their wiring. As these edges are functional in nature, one could either attempt to identify the edges from qualitative data or attempt to infer edges from quantitative data (**Figure 3**). In particular, perturbation experiments in which defined stimuli

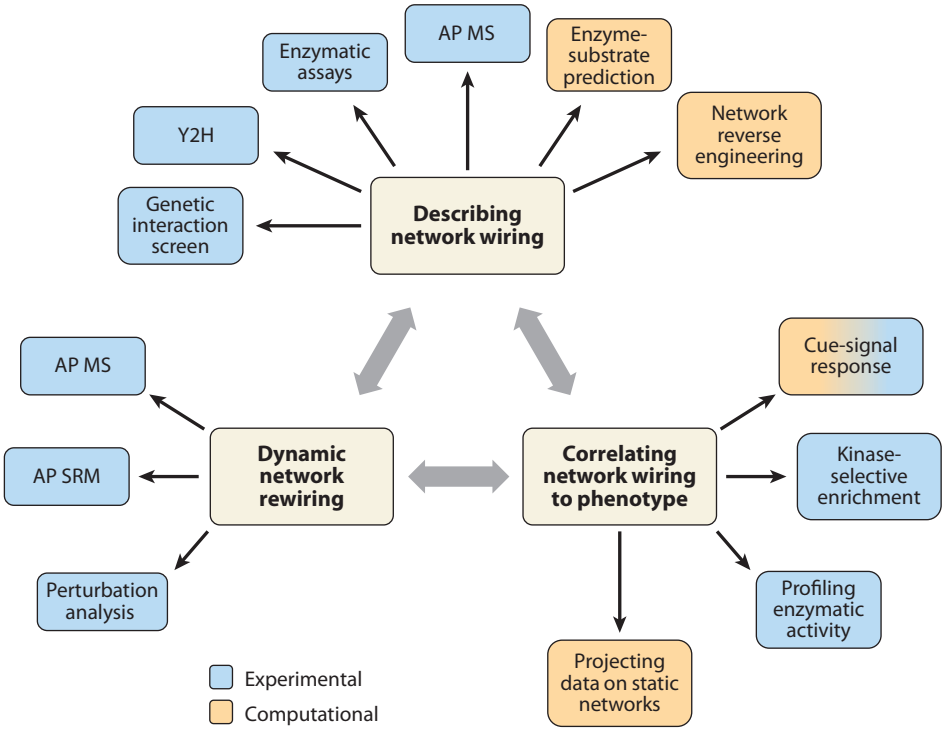


Figure 3

Describing protein-signaling network (PSN) wiring can be performed by either attempting to identify the edges from qualitative data or attempting to infer edges from quantitative data. The yeast two-hybrid (Y2H) system and affinity purification mass spectrometry (AP MS) have been applied to describe protein interaction networks, and various approaches have been applied to assay the relationship between enzymes and substrates. Genetic interaction screens provide complementary information. Dynamic network rewiring in which defined stimuli are used to induce rewiring of the network may highlight the structure of an underlying PSN. It is performed by quantitative analysis of AP-MS or AP-SRM (selected reaction monitoring) experiments and by perturbation experiments using small molecules or gene deletions. Correlating network wiring with phenotypes can use the underlying network of static protein-protein interactions to identify subsets of nodes that have certain network properties and correspond to a given phenotype. Alternatively, the activity of proteins, such as kinases, can be correlated with a given phenotype. In addition, correlating network rewiring with cellular response can be achieved using cue-signal-response compendiums in which samples are perturbed with various cues, preselected network nodes, and cellular phenotypes measured in these samples, and the resulting data are subjected to statistical and modeling frameworks to examine/investigate the signals and responses and to identify those of significance.

are used to induce rewiring of the network may highlight the structure of an underlying PSN. This is discussed further in the next section.

In reviewing PSNs, we refer mainly to protein phosphorylation as a prototypical example of a PSN. Protein phosphorylation is a PTM that has received much attention in proteomics, fueled by rapid improvements in recent years in the enrichment of phosphopeptides from proteome extracts (86–88). More than 160,000 phosphorylation sites are now documented in publicly accessible databases, such as PhosphoSitePlus (89). This vast amount of sites represents the accumulation of data from various studies, and the identification of sites of phosphorylation has become a success story for the application of MS-based proteomic methods used within the molecular biology paradigm.

However, delineating the PSN, i.e., defining which kinases phosphorylate these phosphorylation sites, has lagged behind. Even more challenging have been the questions of which phosphatases dephosphorylate these phosphorylation sites and which other regulatory proteins are involved. Several experimental and computational approaches to provide a description of the kinase-substrate relationships by proteomics have been undertaken (**Figure 3**). They mainly include in vitro kinase assays, prediction of putative kinase-substrate relationships, and perturbation experiments (90, 91).

Methods that are based on in vitro kinase assays provide the most direct data to describe kinase-substrate relationships. For example, Mok et al. (92) conducted protein microarray kinase assays in which a yeast proteome was immobilized on microarrays and incubated with radiolabeled ATP and a kinase of interest. In a conceptually similar approach, peptides were spotted on an array format and phosphorylated by purified kinases to determine consensus recognition motifs or kinases (93). However, the substrates identified in vitro may not represent the bona fide in vivo set of substrates. MS-based proteomics applied to analyze the products of in vitro kinase reactions would obviate the need for radiolabeled ATP and, more importantly, would enable the identification of the phosphorylated residue(s). One successful method relies on engineering a kinase to accept unnatural ATP analogs by modification of the ATP-binding pocket, resulting in the specific thiophosphorylation of a substrate. Substrates of such an engineered kinase are then identified by enrichment of thiophosphopeptides, which can be readily identified by MS (94, 95).

Another approach to capture enzyme-substrate relationships is to stabilize the transient interaction between the two molecules. Bloom et al. (96) expressed in yeast catalytically altered Cdc14 phosphatase mutants that trapped their substrates. Substrates can therefore be isolated while bound to the enzyme and identified by AP MS. "Substrate trapping" has also been applied to a kinase by cross-linking a kinase and a mutated substrate with a specific cross-linker (97).

Finally, a variety of tools have emerged in recent years to predict which kinases putatively phosphorylate a given phosphorylation site, relying on amino acid sequence motifs and a variety of computational methods (91). Of these, NetworKIN (98) should be specifically mentioned because it also uses probabilistic protein association networks. A shortcoming of kinase-substrate predictions is the slower rate at which refinement of kinase consensus motifs occurs, compared to the higher rate at which phosphoproteomics data are acquired. Moreover, these methods do not take into account information from quantitative measurements and rely mostly only on sequence information.

A different computational strategy to infer the wiring of a signaling network is referred to as reconstruction or reverse engineering. Variants of this strategy have been mainly applied to microarray data (reviewed in References 99 and 100). These methods attempt to identify causal relationships between network nodes from quantitative omics data, such as the transcription factors that control modules of coregulated transcripts in the case of microarray data. To the best of our knowledge, MS-based proteomics data have not been used in such approaches, but there is a clear analogy between transcription factors and transcripts measured by microarrays and kinases and the phosphorylation sites measured by phosphoproteomics. However, network inference is not a trivial computational task, and it might not be easily generalized and transferable to MS-based proteomics data.

Other modifications have also been subject to large-scale MS analyses, such as N-terminal and lysine acetylation (41, 101) and N-glycosylation (43). In these studies, thousands of the modified sites have been identified, attesting to the breadth of the identifiable PTM landscape. As in the case of phosphorylation, proteins modified by these PTMs and others are most successfully identified after enrichment. Several studies in recent years have also addressed the identification of

proteins modified by ubiquitin and ubiquitin-like modifications, which are small conserved proteins themselves that covalently modify other proteins in a reversible manner (102, 103). In contrast to the rapid increase in studies focused on detecting dynamic change of the phosphoproteome in perturbed cells, the analysis of ubiquitome thus far has lagged behind and has mainly focused on the identification of modified proteins and sites in steady state. Argenzio et al. (104) have explored the epidermal growth factor (EGF)-regulated ubiquitome by a quantitative analysis of cells untreated or treated with EGF. Kim et al. (105) and Wagner et al. (106) used a novel antibody to monitor changes in the ubiquitome in response to proteasomal inhibition. The cleavage of proteins by proteases has also been investigated as a modification in itself, by probing the activity and specificity of proteases (107, 108), as well as by identification of protease cleavage products (109, 110). As there are hundreds of proteins putatively involved in these various protein modifications, clearly the analysis of the corresponding signaling networks is even further in the future than for protein phosphorylation.

DYNAMIC NETWORK REWIRING

MS-based proteomic techniques have been applied toward identifying and quantifying changes in network wiring in response to stimuli or under different physiological conditions. For the most part, these techniques have been used in the molecular biology paradigm to detect abundance changes in network nodes. In the following section, we review studies that focus on either quantifying changes in known edges or identifying novel perturbation-specific edges, thus describing in detail novel network wiring (mentioned in the section above).

Dynamic Rewiring of Protein Interaction Networks

To date most PIN, particularly PPIN, studies have been performed under basal conditions. They do not, therefore, include information about dynamics or reorganization of these PPINs. Understanding the dynamic nature of PPINs as a function of time and/or condition requires quantification of the proteins that are retained, released, or recruited into a complex in response to a perturbation. Few studies so far have attempted time- or perturbation-resolved dynamics. Examples of these studies include the FoxO3A interactomes in response to phosphatidylinositol 3 kinase inhibition (73), the dynamics of the extracellular signal-regulated kinase 1 interactome in response to the EGF and nerve growth factor (111), and the dynamics of the circadian rhythm gene *frequency* (FRQ) interactome during the course of a day (112). In work closing the gap to translational use of PPIN knowledge, Aye et al. (113) applied a chemical proteomics approach to tissue samples of human patients to demonstrate that the PPIN profile of the regulatory subunit of protein kinase A with several of its scaffold proteins became severely altered in the failing heart.

Several techniques have been applied to resolve changes in complex composition: Wepf et al. (114) used a reference peptide, dubbed SH-quant, as part of a protein's affinity tag to calculate the absolute amounts of bait proteins and, by correlational quantification, the amounts of the prey proteins in various AP samples. This approach was then used for the analysis of perturbation-induced quantitative changes in the composition of the protein phosphatase 2A complex, thus resolving protein-complex abundance and dynamic changes in complex components by combining the labeled peptide with label-free quantification. Bennett et al. (115) mapped the basal cullin-RING ubiquitin ligase network and then investigated the changes in the network in response to deneddylation. Moreover, they used AQUA peptides (116) and MS quantification to determine changes in the subunit occupancy of the cullin-RING ubiquitin ligase network. The use of targeted proteomics can facilitate the validation of larger numbers of interactors at high sensitivity and throughput. For example, Bisson et al. (117) used a method they termed AP-SRM

AQUA: absolute quantification using synthesized peptides, incorporated with stable isotopes and spiked in a known concentration into samples as internal standards

to study the dynamics of PPINs. They first identified novel interactors of growth factor receptor-bound protein 2 (GRB2), a scaffold protein involved in tyrosine kinase receptor signaling. They then generated SRM assays for 90 interactors and used them to provide a quantitative temporal analysis of the GRB2 PPIN following stimulation with EGF, as well as analysis of growth factor-specific GRB2 PPIs. The success of these studies suggests that the PPINs will increasingly be studied as dynamic networks using SRM-based targeted-MS techniques.

Dynamic Rewiring of Protein-Signaling Networks

Perturbation experiments try to capture the effects of modulating an enzyme activity (inhibition or activation) in a cellular context. The dynamic rewiring of the network could then lead to a context-specific description of the network wiring, which is expected to provide new physiological insights compared to in vitro experiments (**Figure 3**). Two conceptual types of perturbations can be pointed out in the context of network biology: those that affect mainly the edge(s) versus those that affect mainly node(s) and inevitably also their edge(s) (118). Perturbation information comes at the cost of increased complexity and treads a fine line between dependency, causality, and the off-target effects of the molecules used.

In one perturbation setup to study phosphorylation networks, a comparison is made between untreated cells, cells treated with a stimulus, and cells treated with a stimulus in the presence of a kinase inhibitor (for example, References 119 and 120). This experiment can be combined with phosphoproteomics, and from such perturbations, phosphopeptides downregulated in the presence of the inhibitor are considered as kinase dependent, although not necessarily directly kinase mediated. In the context of kinase-substrate networks in which kinase specific inhibitors are employed, it is important to remember that inhibitors actually cover a wider range of kinases at different selectivities (121). This is important not only for clinical investigation of these molecules, but also for biologists interested in the interpretation of experimental perturbation data. To circumvent the issue of inhibitor specificity, several groups have adopted the approach of using analog-sensitive kinases, which have been mutated at the ATP-binding pocket to be specifically and rapidly inhibited by the pyrimidine-based inhibitor, 1-NM-PP1 (122). Holt et al. (123) used this approach to identify novel putative substrates of cyclin-dependent kinase 1 (Cdk1) by identifying those phosphorylation sites that conform to Cdk1's known consensus motif and were downregulated upon Cdk1 inhibition. In a few cases, a phosphoantibody is available for the known consensus motif of a kinase and can be used as an affinity reagent to enrich phosphopeptides from samples in a quantitative experimental setup in which the kinase is activated/inhibited (124, 125). Such samples are then compared by quantitative proteomics to suggest substrates for this known kinase.

One should note that dependence does not necessarily indicate that these are true substrates because phosphorylation sites may presumably be regulated by other kinases, which are themselves regulated. Thus, as phosphoproteomics applied in such setups is an indirect measurement of edges that have to be inferred, it may give only a limited scope of the kinase-substrate relationships embedded in the data. However, it can provide insights to the functional organization of a stimulus-dependent phosphoprotein network, e.g., proteins identified in phosphoproteomic screens of perturbed cells are clustered to identify groups of coregulated proteins, and the resulting clusters are projected onto static PPIN data (46, 119, 126).

From a network perspective, inhibition of a kinase by a small and specific molecule is an "edgetic" approach, removing those edges that represent kinase-substrate relationships. Perturbation can also be carried out by removing a node completely, for example, by deletion of the gene. This would affect not only

kinase-substrate relationships, but also kinase interactions with other proteins. Detecting the effects of deleting kinase nodes may be problematic as these effects might be compensated for over time. Bodenmiller et al. (47) analyzed phosphoproteome samples from a large selection of yeast strains in which kinase and phosphatase genes were deleted. The loss of most of these has perturbed significant parts of the phosphoproteome, but only about half of them in the expected direction (e.g., downregulation of a phosphopeptide when a kinase is deleted). These data indicate that the overall wiring of PSNs and their perturbation-induced rewiring is currently beyond the range of MS-based proteomics.

Although these studies tend to be global in approach, two recent studies focused on a specific subset of kinase-substrate relationships and limited their scope of interest in network rewiring to a well-defined biological question. Jorgensen et al. (127) examined bidirectional signaling initiated by cell-cell contacts, studied by the coculture of two cell lines and analysis of their tyrosine phosphoproteomes. The data were used along with results from an siRNA and predicted kinase-substrate networks and PPINs to derive a network model of cell-specific information. Coba et al. (128) examined changes in the phosphoproteome of the neurons' postsynaptic density in response to *N*-methyl-D-aspartate. Combining these data with information on activated kinases, identified by the Western blot method and in vitro kinase assays on peptide arrays, enabled them to formulate a putative kinase-substrate network of the synapse, benefiting from the lower complexity of postsynaptic density compared to a whole cell. Both of these studies pooled descriptive and perturbation-related quantitative data and suggest an inferred wiring of the kinase-substrate relationships. Although successful, both also highlight that MS-based proteomics applied in perturbation setups is still resulting in the description of network wiring and is not rigorously applied to quantify dynamic rewiring per se.

CORRELATING NETWORK WIRING WITH PHENOTYPES

In a network paradigm, we wish to know how networks capture and process information to induce specific cellular responses or phenotypes. Ideally, to correlate specific network structures with phenotypes, subtle rewiring of specific networks would need to be detected and quantified and related to well-defined, quantifiable phenotypes. Both the detection of network changes and the definition of quantitative phenotypes are challenging. It is therefore not surprising that studies in the field are sparse. A significant fraction of the published studies to date attempt to relate clinical phenotypes to changes in cellular networks, particularly in the field of cancer biology.

Correlating Network Wiring with Phenotypes for Protein Interaction Networks

The actual measurement of PPINs and their dynamic change remains technically challenging. This is in contrast to the description of static PPINs, where significant progress has been achieved. Several studies have therefore attempted to use static PPINs as a priori knowledge to investigate correlations in experimental data generated by various technologies. An underlying network, e.g., a static PPIN, is used to identify subsets of genes that have certain network properties and correlate with a phenotype (**Figure 3**). These have been successfully applied in combination with large-scale nucleic acid data. For example, Taylor et al. (129) constructed a large-scale PPIN and then used the average Pearson correlation coefficient to quantify the extent to which a hub and its interacting partners were coexpressed. They investigated a cohort of sporadic, nonfamilial breast cancer patients and detected 256 hubs that displayed altered Pearson correlation coefficients of expression between groups with different clinical outcomes. Mani et al. (130) constructed a network by combining PPI data

and protein-DNA interaction data measured by microarray technology and searched for genes with unusual numbers of edges, which show a change in correlation in a phenotypic subset of the samples. The ranked list can then be presented in the context of the constructed network. Rather than using a global static PPIN, Chang et al. (131) focused on the analysis of confined networks. First, an initial set of genes in a local network was defined, e.g., the interactors of an oncogene, which was then expanded to identify sets of genes whose signatures reflect modular aspects of expression variation. The activity of selected signatures was investigated in a panel of cancer cell lines and tumor samples. In this study, focused on local networks, the novel signatures are anchored to well-known network modules or local networks and may point to novel mechanisms within the initial context. Of note, Nibbe et al. (132) used targets from a proteomic screen as seeds for searching significant subnetworks in colorectal cancer, which included direct interactors as well as "crosstalkers," proteins in the neighborhood of seeds. Subnetworks were then scored, using mRNA expression data, to estimate the significance in differentiating tumor from control samples, allowing data integration regarding dysregulation at both mRNA and protein levels. Finally, in a unique approach, Lage et al. (133) integrated phenotypic data from targeted mutations in mice with PPIN data to describe a functional network underlying cardiac development. This is perhaps a simple example of a switch from the one gene-one protein-one function paradigm toward the network biology paradigm, where molecular networks are viewed as the determinants of phenotypes.

The above summary indicates that, to date, most studies that attempt to correlate molecular networks with phenotypes have combined large-scale, typically static proteomic data with other data resources. The validity of the conclusions derived from correlating quantitative data and network knowledge depends on the coverage and correctness of the network, the quality of the data, and the method of correlation calculation. Furthermore, the results do not necessarily suggest causality.

Correlating Network Wiring with Phenotypes for Protein-Signaling Networks

As for the use of MS-based proteomic data to describe the wiring and rewiring of PSNs (discussed above), most of the literature investigating the connection of PSNs with phenotypes has been focused on protein kinase-substrate networks. The following discussion is therefore also focused on protein phosphorylation but exemplifies other types of PSNs as well.

Protein kinases are attractive drug targets (121). For this reason, it is of interest to correlate the activity of a kinase with a given phenotype. Studies that attempted to unravel kinase-substrate relationships have indicated extensive indirect and compensatory effects, suggesting a complex and yet incompletely understood kinase-substrate network (47; see the Dynamic Network Rewiring section above). These data also suggest that discerning the rationale for pharmacologic inhibition of protein kinases with the aim of affecting disease phenotypes will be most successful in a network biology paradigm.

Several studies have generated activity profiles of kinases in biological samples. Rivoka et al. (134) examined the extent of tyrosine phosphorylation across many carcinoma cell lines and tumors, assuming that the degree of phosphorylation of the kinase correlated with its state of activity. In a more direct approach, Cutillas et al. (135) used MS-based proteomics to quantify a specific phosphopeptide as a surrogate for kinase activity. The peptide in its unphosphorylated form was incubated with a cell lysate and ATP, and after the reaction was quenched, the abundance of the phosphorylated form of the peptide was quantified using a spiked-in internal standard. Kubota et al. (136) applied a similar approach, but with 90 peptides, in a multiplexed manner. Although not every peptide could be assigned to a unique kinase, profiles of activity could be inferred and

were shown to differ between different cell lines and in response to stimuli. In these methods, the signal indicating kinase activity is amplified by the kinase reaction, making even low-abundance kinases detectable, provided they are active. Extending such approaches to a complete kinome level will remain challenging as long as the wiring of the basic kinase-substrate network remains incomplete. Daub et al. (137) therefore profiled the phosphoproteome of kinases themselves by kinase-selective enrichment via affinity chromatography, followed by phosphopeptide enrichment. Monitoring regulatory phosphorylation sites on the enriched kinases, such as phosphorylation events on the activation loop, can then be used as a surrogate marker for changes in kinase activities. The use of kinase inhibitors as affinity reagents is attractive as it can allow capturing dozens of kinases at a time (137, 138). However, the analysis may become complicated in the case of some inhibitors for which the binding affinity of the kinase depends on its activation state and is therefore influenced by the phosphorylation of the activation loop (139).

Given the difficulties of relating complete or extensive PSNs to phenotypes, it is no surprise that studies focused on smaller subnetworks, and selected network nodes have progressed faster (reviewed in References 13, 14, and 140). Such studies are exemplified by the "cue-signal-response" (CSR) compendiums, a type of data acquisition where multiple, typically several dozen, samples are perturbed with various cues—molecules affecting a certain signaling network. Preselected network nodes and cellular phenotypes are then measured in these samples, and the resulting data are subjected to statistical and modeling frameworks to correlate signals with responses. For example, Janes et al. (141) stimulated colon adenocarcinoma cells with nine combinations of tumor necrosis factor, EGF, and insulin (cues), and collected 19 intracellular measurements of the known underlying signaling network (signals) and four apoptotic outputs (responses) along 13 time points. The compendium collected was then analyzed using a data-driven modeling approach to map relationships between the phosphorylation signals measured and cellular death responses. The model captured the two canonical axes of the cellular response, e.g., apoptosis versus survival, and identified previously unknown components of the signaling network, such as autocrine feedback. Saez-Rodriguez et al. (142) applied a different computational approach and used such a compendium to calibrate Boolean logic models of a literature-derived signaling network. They found that a Boolean model, even though it captures only two activation states (on and off), can still fit experimental data and therefore can be used as an approach to harness knowledge already available. The method applied by Janes et al. does not necessarily require a priori mechanistic knowledge (143), but it does entail choosing the informative combinations of CSR, which means a study based on prior knowledge is more likely to provide informative insights. Importantly, although it provides a condensed set of the most informative measurements that fit the data to a model, at the same time the method may overlook key condition-specific modulators, if these conditions are not properly addressed (144). Nelander et al. (145) suggested a method to derive a network structure without any a priori knowledge by applying a CSR setup that includes multiple inputs of drug combinations and measuring multiple outputs of phosphoproteins and phenotypes. Although nodes in the final model are defined as only those precisely perturbed or measured, the model could nevertheless recapitulate known relationships. Finally, Mitsos et al. (146) suggested an approach to identify alterations in a pathway/network of interest in response to drug treatment, regardless of the cellular phenotype. A cell-type-specific network was constructed using integer linear programming and a compendium of phosphorylation site measurements in cells treated with cytokines and known specific inhibitors. The effects of a given drug were detected by reapplying this procedure with the drug instead of the known inhibitor and by identifying altered edges in the model, inferred

as drug specific. This approach does not require the measurement of responses, only cues and signals are measured, but it does require a priori knowledge for the network construction and does not offer a uniquely optimal solution.

What the above studies have in common are datasets of relatively few (typically a dozen or so) selected network nodes that were quantified in multiple perturbed samples and that provided insights to the dynamic rewiring correlating to the phenotype. Interestingly, the majority of these data were generated by quantitative Western blotting, a technique that is limited by the availability of antibodies specific for the selected network nodes. It can be expected that the use of suitable MS-based proteomic techniques for CSR compendiums will derive significant advantages. First, the selection of measured network nodes will no longer be constrained by the availability of antibodies: In principle, every protein or phosphorylation site should be measurable. Second, the number of quantified nodes can be extended with a modest increase in experimental cost. Third, the quantitative accuracy of MS-based proteomics should exceed that of Western blotting, and fourth, the sample throughput should be dramatically increased. However, the advantages MS-based proteomics can offer might cause a dramatic increase in computational cost.

As discussed above (see the Mass Spectrometry–Based Proteomics section), different mass spectrometric strategies differ in their performance profiles. Targeted proteomics by SRM is the method of choice if a limited number of analytes needs to be quantified in multiple samples at a high level of reproducibility, sensitivity, and quantitative accuracy, as is the case in the generation of CSR compendiums. The performance of this approach to study the dynamics of biological networks and the resulting phenotype was demonstrated in two recent studies. Costenoble et al. (45) developed SRM assays for 228 proteins that constitute the central carbon and amino acid metabolic network of yeast. This set of proteins was then quantified at five metabolic states. The data uncovered nutritional environment-dependent changes in protein abundances and identified isoenzymes that are preferentially expressed at specific states. The data also suggested that metabolic proteins, which are, according to current network models, not required for growth under certain nutritional environments, are still expressed, presumably to allow adaptation to rapid changes in environment. In a related study, Picotti et al. (44) quantified a more limited set of metabolic proteins in a series of samples collected across different metabolic shifts. Clustering of the resulting quantitative profiles indicated sets of proteins in the metabolic network that are subject to comparable transcriptional/translational control. These studies demonstrated that the nodes of a known network if selected and targeted can be quantified over the whole dynamic range of protein expression in yeast cells. One of the bottlenecks of SRM is the need to design optimal assays for each protein by using a unique set of peptides and their respective transitions. Recently, experimental and computational approaches have been developed to allow faster and more accurate development of these assays (27, 147–149), and for selected species, databases containing assays for each protein of the respective proteome are being developed. We therefore expect a significant increase in the use of targeted-MS proteomics for CSR compendiums and for the analysis of network-phenotype relationships in general.

OUTLOOK

The network biology paradigm places networks of interacting molecules between genotype and phenotype. It makes the assumptions that the network wiring is dynamically and multidirectionally modulated in response to external and internal stimuli and that such changes determine phenotypic changes. The abilities of MS-based proteomics to describe network wiring, to capture dynamic rewiring of networks in response to stimuli, and to correlate network wiring to phenotypes are constantly advancing.

MS-based proteomics faces major technological challenges in achieving the goals of comprehensive, reproducible, and quantitative description of proteomes at reasonable throughput. As these challenges are addressed, new ones arise. For example, with the rapid increase in throughput, estimating the level of confidence in assignment of MS spectra to peptide sequences is often performed by statistical analysis at a fixed false discovery rate. Over time, this can result in the accumulation of false data if large datasets are combined (150, 151). With the introduction of modeling into proteomics, we also hope to see application of experimental design optimization (152) not only with respect to statistical design issues (153), such as replication and blocking, but also with respect to choosing the informative time points, observables, and measurements.

There is no doubt that the rigorous implementation of network biology will require the development of a range of new (proteomic) technologies that are focused on identifying and quantifying the edges of dynamic networks. Although the description of PPINs has seen much progress, the description of interactions between proteins and other molecules, and the description of the PSN, has lagged behind. For the latter, we expect new experimental setups to address the relationships between enzymes and their substrates for various types of PTMs. Techniques to describe network rewiring in response to stimuli have been only recently emerging, and we still largely lack their application to measuring dynamics of known signaling edges. In this respect, datasets that attempt to measure not only nodes or edges, but also cellular responses to correlate them with, would advance the ability to harness proteomics to network biology.

Ideally, the PPIN and the PSN are evaluated in their genuine biological context. At present, the tools are not yet fully developed to chart the PPIN and the PSN in primary cells, tissues, or whole mammalian organisms. Although the basic wiring of PPINs and PSNs may be highly conserved (even from yeast to human), an extra dimension of complexity will be provided by the tissue-, cell-, and organelle-specific features of PPINs and PSNs. Optimistically, many of the proteomics strategies described here may be transferable to analysis of tissue and even of whole mammals.

Finally, using information already gathered to generate new knowledge is a task performed less often in MS-based proteomics. We would like to see more data being used as the generation of informative datasets is refined. At the moment, we deposit large datasets for wet lab biologists to use, with no knowledge if it is actually employed by them. We thus believe that the transition to network biology approaches and applying modeling will not only provide a new wealth of knowledge but also improve proteomic measurements, moving from "what can we measure" to "what should we measure." An iterative approach in which network-driven biological questions are addressed by MS-based proteomics, combined with computational modeling and used to guide the selection of the next set of experiments, will drive our understanding of biological complexity.

SUMMARY POINTS

1. The new paradigm emerging is typically referred to as systems biology, network biology, or integrated biology. Although molecular biology is concerned with the network nodes involving one-dimensional pathways, the new paradigm entails network nodes and edges and places multidimensional and multidirectional networks of interacting molecules between genotypes and phenotypes.

2. Until recently, mass spectrometry (MS)-based proteomics has been used predominantly to identify, characterize, and quantify network nodes without consideration of their context. While the results from such analyses have increased in volume (longer lists) and confidence, the amount of new biology learned has been moderate.
3. Recently, MS-based proteomics has been used to identify network edges, e.g., describe network wiring. Affinity purification followed by MS has been used to interrogate protein-interaction networks. In vitro assays and perturbation experiments have been used to describe or infer protein-signaling networks (PSNs).
4. Describing the dynamic multidirectional rewiring of networks has been emerging recently in MS-based proteomics. Applications to PPINs include the measurement of dynamic changes in protein complexes, whereas applications to PSNs have mainly focused on perturbation experiments used to infer network wiring.
5. Correlating network wiring to phenotypes using proteomics data has been mostly applied by combining static PPINs with dynamic microarray data. Correlating dynamic proteomic PPIN and PSN data awaits application, and cue-signal-response (CSR) may be an experimental setup for this task.
6. Overall, it is apparent that the MS-based proteomic technologies developed within the molecular biology paradigm are useful but not sufficient for network biology. Conceptually, new technologies, rather than incremental advances of current technologies, will therefore be required.

FUTURE ISSUES

1. Can a comprehensive coverage of a complex proteome be obtained in a reproducible manner, with high quantitative accuracy, and at moderate to high throughput?
2. Given that most data collected to date are under basal conditions, how can AP-MS workflows be improved to allow quantitative dynamics of network rewiring?
3. How can the identification or inference of edges in PSNs and their dynamic changes be improved to be quantitative, reproducible, and comprehensive? Can large-scale dynamic rewiring be captured by MS-based proteomics?
4. How can MS proteomics be applied to collect CSR compendiums? Considering their complexity, how can these be analyzed computationally?
5. How can computational analysis of prior data be used in experimental design to direct optimal MS proteomic measurements?
6. Will MS-based proteomics generate data for mechanistic models of signaling networks?
7. How can emerging proteomics technologies that probe network biology be efficiently transferred to primary cells, tissue, and whole animals?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Work in our laboratory is supported by research grants from the European Community's Seventh Framework Program [the TRIREME Project (grant 223575), the PROSPECTS Project (grant 201648), SYBILLA (grant 201106), SYSTEM TB (grant 241587), and PRIME-XS (grant 262067)], Swiss National Science Foundation (grant 3100A0-107679), SystemsX.ch, and the Swiss Initiative for Systems Biology, as well as by funds from the ERC advanced grant, Proteomics v 3.0 (grant 233226). A.J.R.H. kindly acknowledges support of the Department of Biology of the ETH Zurich, enabling his sabbatical residence at the Institute for Molecular Systems Biology, which was also supported by a Distinguished Visiting Scientist Stipend of the Netherlands Genomics Initiative.

LITERATURE CITED

1. Beadle GW, Tatum EL. 1941. Genetic control of biochemical reactions in *Neurospora*. *Proc. Natl. Acad. Sci. USA* 27:499–506
2. Hawkins RD, Hon GC, Ren B. 2010. Next-generation genomics: an integrative approach. *Nat. Rev. Genet.* 11:476–86
3. Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J. 2010. The next-generation sequencing technology and application. *Protein Cell* 1:520–36
4. Cox J, Mann M. 2011. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu. Rev. Biochem.* 80:273–99
5. Le Provost F, Lillico S, Passet B, Young R, Whitelaw B, Vilotte JL. 2010. Zinc finger nuclease technology heralds a new era in mammalian transgenesis. *Trends Biotechnol.* 28:134–41
6. Boutros M, Ahringer J. 2008. The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.* 9:554–66
7. Zhang EE, Liu AC, Hirota T, Miraglia LJ, Welch G, et al. 2009. A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell* 139:199–210
8. Kondo S, Perrimon N. 2011. A genome-wide RNAi screen identifies core components of the G-M DNA damage checkpoint. *Sci. Signal.* 4:rs1
9. Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, et al. 2009. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* 139:1255–67
10. Cazier JB, Tomlinson I. 2010. General lessons from large-scale studies to identify human cancer predisposition genes. *J. Pathol.* 220:255–62
11. Amberger J, Bocchini C, Hamosh A. 2011. A new face and new challenges for Online Mendelian Inheritance in Man (OMIM®). *Hum. Mutat.* 32:564–67
12. Arkin AP, Schaffer DV. 2011. Network news: innovations in 21st century systems biology. *Cell* 144:844–49
13. Kreeger PK, Lauffenburger DA. 2010. Cancer systems biology: a network modeling perspective. *Carcinogenesis* 31:2–8
14. Pe'er D, Hacohen N. 2011. Principles and strategies for developing network models in cancer. *Cell* 144:864–73
15. Vidal M, Cusick ME, Barabasi AL. 2011. Interactome networks and human disease. *Cell* 144:986–98
16. Yates JR, Ruse CI, Nakorchevsky A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* 11:49–79
17. Domon B, Aebersold R. 2010. Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* 28:710–21
18. Nesvizhskii AI, Vitek O, Aebersold R. 2007. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat. Methods* 4:787–97
19. Mallick P, Kuster B. 2010. Proteomics: a pragmatic perspective. *Nat. Biotechnol.* 28:695–709
20. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, et al. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1:376–86

21. Mueller LN, Rinner O, Schmidt A, Letarte S, Bodenmiller B, et al. 2007. *SuperHirn*—a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* 7:3470–80
22. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, et al. 2011. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11:535–53
23. Jaffe JD, Keshishian H, Chang B, Addona TA, Gillette MA, Carr SA. 2008. Accurate inclusion mass screening: a bridge from unbiased discovery to targeted assay development for biomarker verification. *Mol. Cell. Proteomics* 7:1952–62
24. Schmidt A, Claassen M, Aebersold R. 2009. Directed mass spectrometry: towards hypothesis-driven proteomics. *Curr. Opin. Chem. Biol.* 13:510–17
25. Lange V, Picotti P, Domon B, Aebersold R. 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* 4:222
26. Gallien S, Duriez E, Domon B. 2011. Selected reaction monitoring applied to proteomics. *J. Mass Spectrom.* 46:298–312
27. Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, et al. 2010. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat. Methods* 7:43–46
28. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, et al. 2009. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat. Biotechnol.* 27:633–41
29. Geromanos SJ, Vissers JP, Silva JC, Dorschel CA, Li GZ, et al. 2009. The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* 9:1683–95
30. Geiger T, Cox J, Mann M. 2010. Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell. Proteomics* 9:2252–61
31. Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR. 2004. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* 1:39–45
32. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, et al. 2009. Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean. *Anal. Chem.* 81:6481–88
33. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, et al. 2012. Targeted data extraction of the MS/MS spectra generated by data independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics.* In press
34. Ahrens CH, Brunner E, Qeli E, Basler K, Aebersold R. 2010. Generating and navigating proteome maps using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* 11:789–801
35. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, et al. 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455:1251–54
36. Malmstrom J, Beck M, Schmidt A, Lange V, Deutsch EW, Aebersold R. 2009. Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* 460:762–65
37. Brunner E, Ahrens CH, Mohanty S, Baetschmann H, Loevenich S, et al. 2007. A high-quality catalog of the *Drosophila melanogaster* proteome. *Nat. Biotechnol.* 25:576–83
38. Munoz J, Low TY, Kok YJ, Chin A, Frese CK, et al. 2011. The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. *Mol. Syst. Biol.* 7:550
39. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, et al. 2011. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* 7:548
40. Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, et al. 2011. The quantitative proteome of a human cell line. *Mol. Syst. Biol.* 7:549
41. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, et al. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325:834–40
42. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, et al. 2010. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* 143:1174–89
43. Zielinska DF, Gnad F, Wisniewski JR, Mann M. 2010. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141:897–907
44. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. 2009. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* 138:795–806

45. Costenoble R, Picotti P, Reiter L, Stallmach R, Heinemann M, et al. 2011. Comprehensive quantitative analysis of central carbon and amino-acid metabolism in *Saccharomyces cerevisiae* under multiple conditions by targeted proteomics. *Mol. Syst. Biol.* 7:464
46. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, et al. 2010. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* 3:ra3
47. Bodenmiller B, Wanka S, Kraft C, Urban J, Campbell D, et al. 2010. Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci. Signal.* 3:rs4
48. Van Hoof D, Muñoz J, Braam SR, Pinkse MW, Linding R, et al. 2009. Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell* 5:214–26
49. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, et al. 2011. Global quantification of mammalian gene expression control. *Nature* 473:337–42
50. Schmidt A, Beck M, Malmström J, Lam H, Claassen M, et al. 2011. Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Mol. Syst. Biol.* 7:510
51. Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4:44–57
52. Ma'ayan A. 2008. Network integration and graph analysis in mammalian molecular systems biology. *IET Syst. Biol.* 2:206–21
53. Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, et al. 2010. Visualization of omics data for systems biology. *Nat. Methods* 7:S56–68
54. Park PJ. 2009. ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* 10:669–80
55. Suter B, Kittanakom S, Stagljar I. 2008. Two-hybrid technologies in proteomics research. *Curr. Opin. Biotechnol.* 19:316–23
56. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, et al. 2001. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292:929–34
57. Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, et al. 2008. High-quality binary protein interaction map of the yeast interactome network. *Science* 322:104–10
58. Bandyopadhyay S, Mehta M, Kuo D, Sung MK, Chuang R, et al. 2010. Rewiring of genetic networks in response to DNA damage. *Science* 330:1385–89
59. Butter F, Scheibe M, Morl M, Mann M. 2009. Unbiased RNA-protein interaction screen by quantitative proteomics. *Proc. Natl. Acad. Sci. USA* 106:10626–31
60. Bantscheff M, Scholten A, Heck AJ. 2009. Revealing promiscuous drug-target interactions by chemical proteomics. *Drug Discov. Today* 14:1021–29
61. Rix U, Superti-Furga G. 2009. Target profiling of small molecules by chemical proteomics. *Nat. Chem. Biol.* 5:616–24
62. Li X, Gianoulis TA, Yip KY, Gerstein M, Snyder M. 2010. Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses. *Cell* 143:639–50
63. Gingras AC, Gstaiger M, Raught B, Aebersold R. 2007. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* 8:645–54
64. Sharon M, Robinson CV. 2007. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annu. Rev. Biochem.* 76:167–93
65. Heck AJ. 2008. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* 5:927–33
66. Leitner A, Walzthoeni T, Kahraman A, Herzog F, Rinner O, et al. 2010. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* 9:1634–49
67. Sigal A, Danon T, Cohen A, Milo R, Geva-Zatorsky N, et al. 2007. Generation of a fluorescently labeled endogenous protein library in living human cells. *Nat. Protoc.* 2:1515–27
68. de Boer E, Rodriguez P, Bonte E, Krijgsveld J, Katsantoni E, et al. 2003. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. USA* 100:7480–85
69. Hubner NC, Bird AW, Cox J, Splettstoesser B, Bandilla P, et al. 2010. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* 189:739–54

70. Malovannaya A, Lanz RB, Jung SY, Bulynko Y, Le NT, et al. 2011. Analysis of the human endogenous coregulator complexome. *Cell* 145:787–99
71. Boulon S, Ahmad Y, Trinkle-Mulcahy L, Verheggen C, Cobley A, et al. 2010. Establishment of a protein frequency library and its application in the reliable identification of specific protein interaction partners. *Mol. Cell. Proteomics* 9:861–79
72. Trinkle-Mulcahy L, Boulon S, Lam YW, Urcia R, Boisvert FM, et al. 2008. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* 183:223–39
73. Rinner O, Mueller LN, Hubálek M, Müller M, Gstaiger M, Aebersold R. 2007. An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. *Nat. Biotechnol.* 25:345–52
74. Glatter T, Wepf A, Aebersold R, Gstaiger M. 2009. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol. Syst. Biol.* 5:237
75. Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, et al. 2008. Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc. Natl. Acad. Sci. USA* 105:1454–59
76. Sowa ME, Bennett EJ, Gygi SP, Harper JW. 2009. Defining the human deubiquitinating enzyme interaction landscape. *Cell* 138:389–403
77. Choi H, Larsen B, Lin ZY, Breitkreutz A, Mellacheruvu D, et al. 2011. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat. Methods* 8:70–73
78. Behrends C, Sowa ME, Gygi SP, Harper JW. 2010. Network organization of the human autophagy system. *Nature* 466:68–76
79. Kuhner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, et al. 2009. Proteome organization in a genome-reduced bacterium. *Science* 326:1235–40
80. Breitkreutz A, Choi H, Sharom JR, Boucher L, Neduva V, et al. 2010. A global protein kinase and phosphatase interaction network in yeast. *Science* 328:1043–46
81. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, et al. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142:967–80
82. Lee KK, Sardiu ME, Swanson SK, Gilmore JM, Torok M, et al. 2011. Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Mol. Syst. Biol.* 7:503
83. Wang J, Li M, Deng Y, Pan Y. 2010. Recent advances in clustering methods for protein interaction networks. *BMC Genomics* 11(Suppl. 3):S10
84. Choi H, Kim S, Gingras AC, Nesvizhskii AI. 2010. Analysis of protein complexes through model-based biclustering of label-free quantitative AP-MS data. *Mol. Syst. Biol.* 6:385
85. Scott JD, Pawson T. 2009. Cell signaling in space and time: where proteins come together and when they're apart. *Science* 326:1220–24
86. Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. 2007. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* 4:231–37
87. Lemeer S, Heck AJ. 2009. The phosphoproteomics data explosion. *Curr. Opin. Chem. Biol.* 13:414–20
88. Thingholm TE, Jensen ON, Larsen MR. 2009. Analytical strategies for phosphoproteomics. *Proteomics* 9:1451–68
89. Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, Zhang B. 2004. PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* 4:1551–61
90. Cutillas PR, Jørgensen C. 2011. Biological signalling activity measurements using mass spectrometry. *Biochem. J.* 434:189–99
91. Tan CS, Linding R. 2009. Experimental and computational tools useful for (re)construction of dynamic kinase-substrate networks. *Proteomics* 9:5233–42
92. Mok J, Im H, Snyder M. 2009. Global identification of protein kinase substrates by protein microarray analysis. *Nat. Protoc.* 4:1820–27
93. Mok J, Kim PM, Lam HY, Piccirillo S, Zhou X, et al. 2010. Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci. Signal.* 3:ra12
94. Allen JJ, Li M, Brinkworth CS, Paulson JL, Wang D, et al. 2007. A semisynthetic epitope for kinase substrates. *Nat. Methods* 4:511–16

95. Blethrow JD, Glavy JS, Morgan DO, Shokat KM. 2008. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proc. Natl. Acad. Sci. USA* 105:1442–47
96. Bloom J, Cristea IM, Procko AL, Lubkov V, Chait BT, et al. 2011. Global analysis of Cdc14 phosphatase reveals diverse roles in mitotic processes. *J. Biol. Chem.* 286:5434–45
97. Maly DJ, Allen JA, Shokat KM. 2004. A mechanism-based cross-linker for the identification of kinase-substrate pairs. *J. Am. Chem. Soc.* 126:9160–61
98. Linding R, Jensen LJ, Pasculescu A, Olhovsky M, Colwill K, et al. 2008. NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* 36:D695–99
99. Karlebach G, Shamir R. 2008. Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* 9:770–80
100. De Smet R, Marchal K. 2010. Advantages and limitations of current network inference methods. *Nat. Rev. Microbiol.* 8:717–29
101. Helbig AO, Gauci S, Raijmakers R, van Breukelen B, Slijper M, et al. 2010. Profiling of N-acetylated protein termini provides in-depth insights into the N-terminal nature of the proteome. *Mol. Cell. Proteomics* 9:928–39
102. Kerscher O, Felberbaum R, Hochstrasser M. 2006. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu. Rev. Cell Dev. Biol.* 22:159–80
103. Shi Y, Xu P, Qin J. 2011. Ubiquitinated proteome: ready for global? *Mol. Cell. Proteomics* 10:R110.006882
104. Argenzio E, Bange T, Oldrini B, Bianchi F, Peesari R, et al. 2011. Proteomic snapshot of the EGF-induced ubiquitin network. *Mol. Syst. Biol.* 7:462
105. Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, et al. 2011. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* 44:325–40
106. Wagner SA, Beli P, Weinert BT, Nielsen ML, Cox J, et al. 2011. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol. Cell. Proteomics* 10:M111 013284
107. Schilling O, auf dem Keller U, Overall CM. 2011. Protease specificity profiling by tandem mass spectrometry using proteome-derived peptide libraries. *Methods Mol. Biol.* 753:257–72
108. Nomura DK, Dix MM, Cravatt BF. 2010. Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat. Rev. Cancer* 10:630–38
109. Mahrus S, Trinidad JC, Barkan DT, Sali A, Burlingame AL, Wells JA. 2008. Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell* 134:866–76
110. Kleifeld O, Doucet A, auf dem Keller U, Prudova A, Schilling O, et al. 2010. Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat. Biotechnol.* 28:281–88
111. von Kriegsheim A, Baiocchi D, Birtwistle M, Sumpton D, Bienvenut W, et al. 2009. Cell fate decisions are specified by the dynamic ERK interactome. *Nat. Cell Biol.* 11:1458–64
112. Baker CL, Kettenbach AN, Loros JJ, Gerber SA, Dunlap JC. 2009. Quantitative proteomics reveals a dynamic interactome and phase-specific phosphorylation in the *Neurospora* circadian clock. *Mol. Cell* 34:354–63
113. Aye TT, Soni S, van Veen TA, van der Heyden MA, Cappadona S, et al. 2012. Reorganized PKA-AKAP associations in the failing human heart. *J. Mol. Cell. Cardiol.* 52:511–18
114. Wepf A, Glatter T, Schmidt A, Aebersold R, Gstaiger M. 2009. Quantitative interaction proteomics using mass spectrometry. *Nat. Methods* 6:203–5
115. Bennett EJ, Rush J, Gygi SP, Harper JW. 2010. Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell* 143:951–65
116. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* 100:6940–45
117. Bisson N, James DA, Ivosev G, Tate SA, Bonner R, et al. 2011. Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. *Nat. Biotechnol.* 29:653–58
118. Zhong Q, Simonis N, Li QR, Charloteaux B, Heuze F, et al. 2009. Edgetic perturbation models of human inherited disorders. *Mol. Syst. Biol.* 5:321
119. Bensimon A, Schmidt A, Ziv Y, Elkon R, Wang SY, et al. 2010. ATM-dependent and -independent dynamics of the nuclear phosphoproteome after DNA damage. *Sci. Signal.* 3:rs3

120. Yu Y, Yoon SO, Poulogiannis G, Yang Q, Ma XM, et al. 2011. Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. *Science* 332:1322–26
121. Karaman MW, Herrgard S, Treiber DK, Gallant P, Atteridge CE, et al. 2008. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 26:127–32
122. Koch A, Hauf S. 2010. Strategies for the identification of kinase substrates using analog-sensitive kinases. *Eur. J. Cell Biol.* 89:184–93
123. Holt LJ, Tuch BB, Villen J, Johnson AD, Gygi SP, Morgan DO. 2009. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* 325:1682–86
124. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER 3rd, Hurov KE, et al. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316:1160–66
125. Moritz A, Li Y, Guo A, Villen J, Wang Y, et al. 2010. Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Sci. Signal.* 3:ra64
126. Mayya V, Lundgren DH, Hwang SI, Rezaul K, Wu L, et al. 2009. Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. *Sci. Signal.* 2:ra46
127. Jørgensen C, Sherman A, Chen GI, Pasculescu A, Poliakov A, et al. 2009. Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science* 326:1502–9
128. Coba MP, Pocklington AJ, Collins MO, Kopanitsa MV, Uren RT, et al. 2009. Neurotransmitters drive combinatorial multistate postsynaptic density networks. *Sci. Signal.* 2:ra19
129. Taylor IW, Linding R, Warde-Farley D, Liu Y, Pesquita C, et al. 2009. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat. Biotechnol.* 27:199–204
130. Mani KM, Lefebvre C, Wang K, Lim WK, Basso K, et al. 2008. A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Mol. Syst. Biol.* 4:169
131. Chang JT, Carvalho C, Mori S, Bild AH, Gatza ML, et al. 2009. A genomic strategy to elucidate modules of oncogenic pathway signaling networks. *Mol. Cell* 34:104–14
132. Nibbe RK, Koyutürk M, Chance MR. 2010. An integrative -omics approach to identify functional sub-networks in human colorectal cancer. *PLoS Comput. Biol.* 6:e1000639
133. Lage K, Mollgard K, Greenway S, Wakimoto H, Gorham JM, et al. 2010. Dissecting spatio-temporal protein networks driving human heart development and related disorders. *Mol. Syst. Biol.* 6:381
134. Rikova K, Guo A, Zeng Q, Possemato A, Yu J, et al. 2007. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131:1190–203
135. Cutillas PR, Khwaja A, Graupera M, Pearce W, Gharbi S, et al. 2006. Ultrasensitive and absolute quantification of the phosphoinositide 3-kinase/Akt signal transduction pathway by mass spectrometry. *Proc. Natl. Acad. Sci. USA* 103:8959–64
136. Kubota K, Anjum R, Yu Y, Kunz RC, Andersen JN, et al. 2009. Sensitive multiplexed analysis of kinase activities and activity-based kinase identification. *Nat. Biotechnol.* 27:933–40
137. Daub H, Olsen JV, Bairlein M, Gnad F, Oppermann FS, et al. 2008. Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol. Cell* 31:438–48
138. Bantscheff M, Eberhard D, Abraham Y, Bastuck S, Boesche M, et al. 2007. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* 25:1035–44
139. Wodicka LM, Ciceri P, Davis MI, Hunt JP, Floyd M, et al. 2010. Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry. *Chem. Biol.* 17:1241–49
140. Janes KA, Yaffe MB. 2006. Data-driven modelling of signal-transduction networks. *Nat. Rev. Mol. Cell Biol.* 7:820–28
141. Janes KA, Albeck JG, Gaudet S, Sorger PK, Lauffenburger DA, Yaffe MB. 2005. A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. *Science* 310:1646–53
142. Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, et al. 2009. Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Mol. Syst. Biol.* 5:331
143. Cosgrove BD, Alexopoulos LG, Hang TC, Hendriks BS, Sorger PK, et al. 2010. Cytokine-associated drug toxicity in human hepatocytes is associated with signaling network dysregulation. *Mol. BioSyst.* 6:1195–206

144. Wu Y, Johnson GL, Gomez SM. 2008. Data-driven modeling of cellular stimulation, signaling and output response in RAW 264.7 cells. *J. Mol. Signal.* 3:11
145. Nelander S, Wang W, Nilsson B, She QB, Pratilas C, et al. 2008. Models from experiments: combinatorial drug perturbations of cancer cells. *Mol. Syst. Biol.* 4:216
146. Mitsos A, Melas IN, Siminelakis P, Chairakaki AD, Saez-Rodriguez J, Alexopoulos LG. 2009. Identifying drug effects via pathway alterations using an integer linear programming optimization formulation on phosphoproteomic data. *PLoS Comput. Biol.* 5:e1000591
147. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, et al. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966–68
148. Brusniak MY, Kwok ST, Christiansen M, Campbell D, Reiter L, et al. 2011. ATAQS: a computational software tool for high throughput transition optimization and validation for selected reaction monitoring mass spectrometry. *BMC Bioinform.* 12:78
149. Reiter L, Rinner O, Picotti P, Huttenhain R, Beck M, et al. 2011. mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* 8:430–35
150. White FM. 2011. The potential cost of high-throughput proteomics. *Sci. Signal.* 4:pe8
151. Reiter L, Claassen M, Schrimpf SP, Jovanovic M, Schmidt A, et al. 2009. Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry. *Mol. Cell. Proteomics* 8:2405–17
152. Kreutz C, Timmer J. 2009. Systems biology: experimental design. *FEBS J.* 276:923–42
153. Oberg AL, Vitek O. 2009. Statistical design of quantitative mass spectrometry–based proteomic experiments. *J. Proteome Res.* 8:2144–56

Annual Review of Biochemistry
Volume 81, 2012

Contents

Recent Advances in Biochemistry

Indexes

Errata

An online log of corrections to *Annual Review of Biochemistry* articles may be found at http://biochem.annualreviews.org/errata.shtml