

REVIEW

Protein-protein interaction networks: probing disease mechanisms using model systems

Uros Kuzmanov and Andrew Emili*

Abstract

Protein-protein interactions (PPIs) and multi-protein complexes perform central roles in the cellular systems of all living organisms. In humans, disruptions of the normal patterns of PPIs and protein complexes can be causative or indicative of a disease state. Recent developments in the biological applications of mass spectrometry (MS)-based proteomics have expanded the horizon for the application of systematic large-scale mapping of physical interactions to probe disease mechanisms. In this review, we examine the application of MS-based approaches for the experimental analysis of PPI networks and protein complexes, focusing on the different model systems (including human cells) used to study the molecular basis of common diseases such as cancer, cardiomyopathies, diabetes, microbial infections, and genetic and neurodegenerative disorders.

Introduction

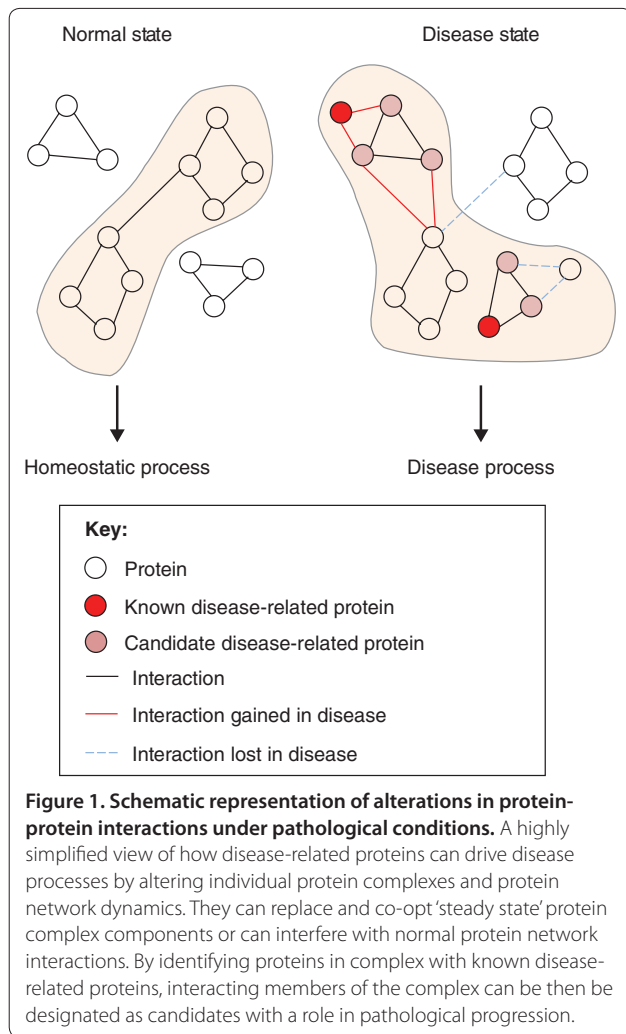
Protein-protein interactions (PPIs) are central to the proper functioning of the most basic molecular mechanisms underlying cellular life, and are often perturbed in disease states. It is predicted that the human complement of PPIs (the interactome) numbers between 130,000 and 600,000 [1,2]. These include interactions of structural proteins inside the cell, and multi-protein complexes that are involved in core processes such as transcription and translation, cell-cell adhesion and communication, protein synthesis and degradation, cell cycle control and signaling cascades. The study of PPI networks and the global physical organization of cells is needed to provide a better understanding of basic cellular biochemistry and physiology (Figure 1). It is therefore no surprise that

when the homeostatic state of an organism or an individual cell is disturbed (as a result of environmental stress or in a disease state) the 'normal' patterns of PPIs are disturbed.

Many of these disruptions can often be considered side products of a disease that have no significant functional consequence, but others can frequently play a major causal role in disease and have a central impact on the initiation or progression of a pathology (Figure 1). For example, the role of PPI disturbances in the interactome of the p53 tumor suppressor protein, caused by mutations in its gene, are well established [3,4]; disruptions in the desmosome-mediated interactions between cells have been implicated in a variety of diseases [5]; aberrant PPIs causing the accumulation of protein aggregates can result in a number of neurodegenerative diseases [6,7]; and host-pathogen PPIs are of central importance in infection [8,9]. Therefore, depending on the pathological scenario, the monitoring and study of PPIs in different biological models can provide interesting and significant options for both diagnostic and therapeutic targets that have potential for broader clinical applicability. The major biomedical goal of identifying and studying PPI networks in disease states is the development of therapies targeting interactions that are functionally relevant to disease progression and patient outcomes. Another long-term clinical goal would be the identification of disease-specific patterns of PPIs, which could serve as disease- or treatment-responsive biomarkers whose selective measurement leads to improved diagnostics or prognostics for common human disorders.

Technological advances in genomics and proteomics have spawned a large number of comprehensive studies that, in turn, have generated huge amounts of data. In recent years, innovative developments in the application of highly sensitive and accurate forms of mass spectrometry (MS) to biological specimens have provided considerable progress in the rapidly emerging fields of metabolomics, lipidomics, glycomics and proteomics. These include the large-scale identification and characterization of a number of post-translational modifications (PTMs) on proteins (phosphorylation, glycosylation, ubiquitylation, methylation and so on). Most notable,

*Correspondence: andrew.emili@utoronto.ca
Banting and Best Department of Medical Research and Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College Street, Toronto, Ontario M5S 3E1, Canada



however, advances in large-scale protein-interaction mapping have led to a significant expansion in our understanding of both the composition of protein complexes and their arrangement within broader cellular PPI networks that are often perturbed under disease states. There have been several reviews of technical developments in the identification and characterization of PPIs and protein complexes [10-13]. Here, we examine the application of MS-based experimental analyses of model systems to explore heterogeneous PPI networks and protein complexes in the context of human disease.

MS-driven interactome studies now serve as a complement to, and extension of, high-throughput mRNA expression profiling and next-generation sequencing platforms. In addition to two-hybrid assay systems, which have been used with great success in mapping individual PPIs, including transient interactions [14-16], MS-based methodologies have become the major tool for the detection of stably co-purifying multi-component (heteromeric) protein complexes. Together, these two tools have

led to the characterization of global PPI networks. In the absence of suitably stringent computational filtering, however, unbiased interaction screens often come at the price of a high false-discovery rate, which necessitates independent experimental validation to verify predicted PPIs.

There are several different types of methodology that utilize MS for the purposes of systematic PPI discovery and global characterization of the components of stable protein complexes. For example, protein complexes can be isolated using affinity purification (AP), using either a tagged 'bait' protein or co-immunoprecipitation (co-IP) if an antibody is available. This is normally followed by 'bottom up' proteomic identification of the purified proteins, which entails proteolytic cleavage of the protein mixture (usually by trypsin) followed by MS-based sequencing of the resulting peptides, from which the protein identities can be deduced. A general workflow for the biochemical isolation of protein complexes and their subsequent MS-based identification is shown in Figure 2. When experimental parameters are optimized, AP/MS-based approaches can often reliably detect interactions for even low-abundance proteins [17], but scaling up to hundreds of targets or more remains a challenge. Conversely, traditional biochemical or chromatographic co-fractionation of endogenous protein complexes has recently been shown to be a viable option for the global profiling of native PPI interaction networks in cell lines (Figure 2), albeit at the cost of reduced sensitivity.

In addition to traditional 'bottom-up' shotgun proteomics-based protein identification, emerging 'targeted' and 'data-independent' acquisition (DIA) MS strategies can also be utilized to monitor PPIs. For DIA MS methods, such as SWATH™ [18], protein identification is achieved by selecting precursor ions for MS2 fragmentation using an incremental mass range window, as opposed to choosing only the most abundant species as during shotgun MS2 sequencing. Conversely, targeted MS approaches, such as selected reaction monitoring (SRM)-based methods (reviewed in [19]), require *a priori* knowledge of the protein components of interest to be analyzed, and hence can only be used to measure pre-selected proteins. Protein-interaction dynamics can be monitored using quantitative MS-based procedures, again in either a targeted or global proteomic fashion. Accurate MS-based global (whole proteome) quantification can be achieved using label-based (for example, stable isotope) or label-free approaches [20].

As far as the biomedical and translational medicine fields are concerned, the major motivation and hope is that the study of PPI networks and protein complexes will yield practical advances in the understanding of the molecular basis of disease processes, which in turn can lead to improvements in diagnostics and therapeutics.

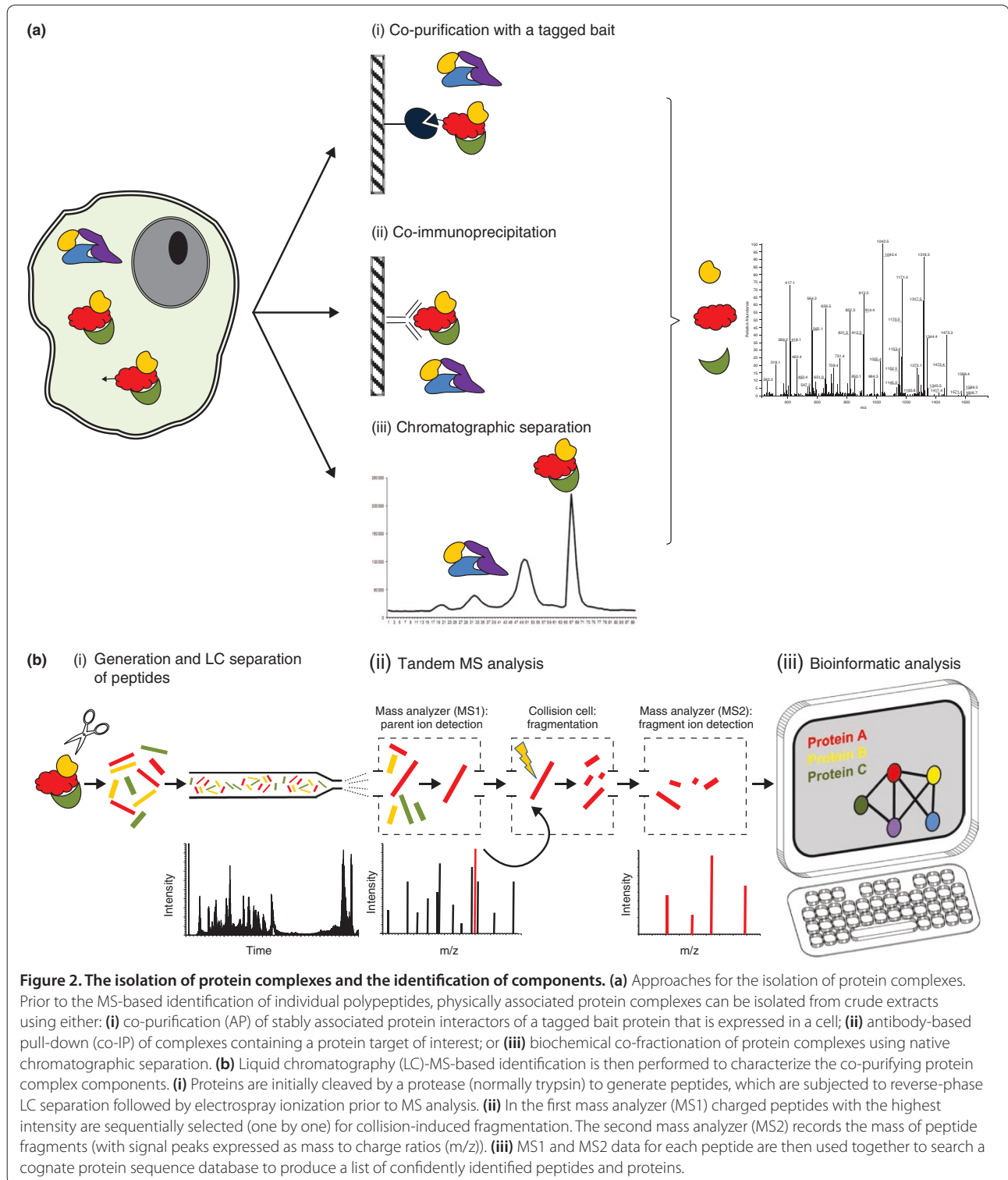


Figure 2. The isolation of protein complexes and the identification of components. (a) Approaches for the isolation of protein complexes. Prior to the MS-based identification of individual polypeptides, physically associated protein complexes can be isolated from crude extracts using either: **(i)** co-purification (AP) of stably associated protein interactors of a tagged bait protein that is expressed in a cell; **(ii)** antibody-based pull-down (co-IP) of complexes containing a protein target of interest; or **(iii)** biochemical co-fractionation of protein complexes using native chromatographic separation. **(b)** Liquid chromatography (LC)-MS-based identification is then performed to characterize the co-purifying protein complex components. **(i)** Proteins are initially cleaved by a protease (normally trypsin) to generate peptides, which are subjected to reverse-phase LC separation followed by electrospray ionization prior to MS analysis. **(ii)** In the first mass analyzer (MS1) charged peptides with the highest intensity are sequentially selected (one by one) for collision-induced fragmentation. The second mass analyzer (MS2) records the mass of peptide fragments (with signal peaks expressed as mass to charge ratios (m/z)). **(iii)** MS1 and MS2 data for each peptide are then used together to search a cognate protein sequence database to produce a list of confidently identified peptides and proteins.

For this goal to be achievable, the above-mentioned methodologies must be applied in the proper context. This is where the choice of model system for any particular disease and the interpretation of the resulting data become crucial. In choosing pertinent studies to address

in this review, we have narrowed the scope by focusing on studies that derive PPIs primarily on the basis of direct experimental data rather than by inference from bioinformatic analysis alone, although some major studies of this latter type will be addressed. Recent studies utilizing

MS for the identification of PPIs relevant in human disease are summarized in Table 1.

Microbes as cell models

Unicellular organisms such as yeast have served as tractable models to probe the molecular biology of eukaryotes, whereas most major human pathogens are prokaryotes. Hence, PPIs have been studied in microbes in great detail. Several landmark studies have contributed greatly to our understanding of the role PPI networks play at all levels of life. The first studies utilizing MS-based approaches in studying PPIs were performed in two of the most basic model systems used in molecular biology, the Gram-negative bacterium *Escherichia coli* and the budding yeast *Saccharomyces cerevisiae*. Owing to their experimental amenability (in terms of genetic manipulation, generation time and so on), these model systems have proven invaluable in proof-of-concept method development in the MS-based interactomics field. Important from a clinical perspective, a significant number of complexes and PPIs that have been mapped in microbes are conserved (to a varying degree) in humans, and disturbances in their normal homeostatic patterns can be indicative or even causative in disease conditions.

The most suitable methodology for study of protein complexes and PPI in these model systems has proven to be the affinity purification of protein complexes followed by MS identification (AP-MS). The existence of genome-scale libraries of genetically engineered *E. coli* and yeast strains expressing individually tagged proteins from native promoters has allowed for the relatively rapid isolation and large-scale mapping of stable protein interactomes in both of these organisms, including most recently membrane-associated complexes [21]. Tandem affinity purification (TAP) [22,23] and sequential peptide affinity (SPA) tagging technologies [24,25] have also contributed to the streamlining of AP-MS identification and characterization of PPIs and heterogeneous protein complexes. These methods allowed for the unprecedented characterization of widely conserved protein complexes in yeast [26] and *E. coli* [27].

Because they are eukaryotic and show a greater degree of conservation with humans, baker's yeast has been a particularly informative model of human protein complexes and PPIs. Several landmark studies have utilized AP-MS to map the yeast protein interactome in a comprehensive manner [28-33]. Two of the more comprehensive studies, from our group and that of a competing company (Cellzome), applied matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) and liquid chromatography (LC)-MS in association with TAP of more than 4,500 tagged yeast proteins to map more than 7,000 interactions and to identify 429 putative protein complexes [26,34]. Notable aspects of the two

studies were the high technical reproducibility and the reciprocal tagging and purification of candidate interactors that provided an estimate of reliability. Strikingly, however, despite using a similarly stringent experimental approach and being co-published at the same time, the overlap of the predicted complexes and PPI was initially found to be low. This discrepancy was widely interpreted as suggesting the incompleteness or unreliability of high-throughput interaction data, but it was later ascribed to differences in the computational scoring and post-processing of each PPI network, indicating that inconsistent data analysis is a major outstanding issue for the field. In a more recent follow-up study in yeast by our group, a carefully defined set of 501 heterogeneous membrane protein complexes were charted in yeast through the additional analysis and identification of detergent-solubilized proteins [21]. A protein kinase-phosphatase interaction network encompassing transient dynamic regulator-substrate interactions has also been mapped using a modified AP-MS-based approach [35].

Owing to the requirement for novel therapeutics and the related need for understanding molecular pathogenesis, PPIs involving pathogenic bacteria and viruses have also garnered significant attention. In the study of viruses, the major focus is the discovery of novel protein-based antigens for the development of vaccines. The mechanisms of host-pathogen interactions and how the pathogen co-opts the host's molecular machinery have also been studied through the examination of host-pathogen PPIs [8]. MS-based methodologies for virus-host proteomics have been reviewed recently [9]. As a result of recent studies of the HIV interactome, several host and viral proteins have been discovered to play a crucial role in the life cycle of infection and appear to have provided potential novel therapeutic targets. An extensive AP-MS-based study of the HIV host-pathogen PPIs was performed [36] by expression of individual tagged HIV proteins transiently in the human embryonic kidney 293 (HEK293) cell line or stably in Jurkat cells (immortalized T lymphocytes) [37]. Putative PPIs from AP-MS were confirmed by co-expression of the strep-tagged viral protein and the TAP-tagged host proteins predicted to interact with it, followed by MS and western blot validation. Using this approach, all 18 HIV-1 proteins were shown with high confidence to be involved in 497 PPIs together with 435 host proteins [36]. A mixture of approaches, including tag-based AP and co-IP followed by MS identification, has been used to identify the host proteins that interact with the HIV pre-integration complex, a key nucleoprotein required for the insertion of the reverse-transcribed viral DNA [38]. MS-based experiments were performed using infected CD4+ human cells.

Other recent examples of viral PPI proteomics studies include the identification of 579 host (human) proteins

Table 1. Recent studies utilizing MS for identification of PPIs relevant in human disease

Model system	Disease	Approach	Reference(s)
Microbes	HIV infection	AP-MS, co-IP-MS identification of HIV-host PPIs	[36-38]
	Viral infection	AP-MS identification of virus-host PPIs	[39]
	Measles	AP-MS identification of virus-host PPIs	[40]
	Hepatitis C	AP-MS identification of virus-host PPIs	[41]
	Japanese encephalitis	AP-MS identification of virus-host PPIs	[42]
	Lyme disease	Biochemical fractionation followed by native gel MS identification of protein complexes in outer membrane of <i>Borrelia burgdorferi</i>	[45]
	<i>Chlamydia</i> infection	Shotgun proteomic identification of <i>Chlamydia trachomatis</i> outer membrane protein complexes	[46]
	Bacterial meningitis	Two-dimensional native gel MS identification of outer membrane protein complexes in <i>Neisseria meningitidis</i>	[47]
	MRSA infection	AP-MS identification of <i>Staphylococcus aureus</i> protein complexes	[48]
	Pneumonia	AP-MS of protein complexes in <i>Mycoplasma pneumoniae</i>	[49]
Mouse	Multiple sclerosis	AP-MS identification of LRP-1-interacting proteins	[56]
	Huntington's disease	AP-MS identification of huntingtin-interacting proteins	[57,58]
	Bovine spongiform encephalopathy	AP-MS identification of prion protein interactors	[60]
	Schizophrenia	AP-MS identification of DLG4 protein interactors	[59]
Human	Breast cancer	AP-MS identification of SCRIB and ER alpha-interacting proteins	[71]
	Leber congenital amaurosis	AP-MS identification of proteins interacting with mutant lebercilin	[72]
	Muscular dystrophy	Co-IP MS identification of dystrophin-interacting proteins	[73]
	Lung cancer	AP-MS identification of EGFR-associated proteins	[75]
	Melanoma	AP-MS identification of HIF2-interacting proteins	[76]

MRSA, methicillin-resistant *Staphylococcus aureus*.

interacting with 70 open reading frames from 30 different viral species. This work utilized TAP-MS to shed new light on conserved viral mechanisms that disrupt host molecular mechanisms [39]. A pilot study examining the PPIs of the tagged MV-V protein (an important virulence factor) from the measles virus utilized AP-MS to find interactions with proteins found in an infected host cell [40]. Identification of the protein-based interactors of the hepatitis C virus protein NS3/4A (which has several roles essential for interaction with host cells) resulted in the discovery of a host protein (Y-box-binding protein 1) that is crucial to the lifecycle of this virus [41]. By identifying the host's binding partners that interact with the core proteins of the Japanese encephalitis virus (a mosquito-borne pathogen), insights were gained into how this pathogen co-opts the host's cellular machinery to ensure propagation [42].

Escherichia coli has proven to be an ideal model system for the study of interaction networks in bacteria. A global map of close to 6,000 PPIs in *E. coli* covering hundreds of protein products of previously uncharacterized 'orphan' bacterial genes was recently published by our group [43]. This study utilized AP-MS to identify binding partners of tagged unannotated proteins, which allowed for their

functional classification following integration with existing genomic data, and revealed many unexpected and diverse functional associations. In a rare example of a non-AP-based approach, 30 *E. coli* putative membrane-associated protein complexes were also identified using a combination of subcellular fractionation with extensive ion exchange chromatography followed by MS identification of co-eluting polypeptides [44].

The direct examination of PPIs in pathogenic bacteria, either in interactions with the host or within the microbe itself, has also attracted some attention. Protein complexes in bacterial membranes have particular relevance both for antigen identification, which can be used for the generation of vaccines, and because of the presence of integral antibiotic clearing pumps. For example, the outer membrane vesicle protein complexes of the Lyme disease parasite *Borrelia burgdorferi* were recently identified [45]. A shotgun proteomic comparison of different subcellular fractions and subsequent bioinformatic analysis allowed the identification of outer membrane complexes of *Chlamydia trachomatis*, providing insights into this bacterium's protein-secretion processes and infectious particle composition, which might be useful for future therapies [46]. Likewise, the outer membrane protein

complexes of *Neisseria meningitidis* (the pathogen responsible for a number of meningococcal diseases) were also recently elucidated using two-dimensional native gel electrophoresis of intact macromolecules followed by MS [47]. Perhaps most impressively, a PPI map of 608 proteins present in methicillin-resistant *Staphylococcus aureus* (a potentially lethal bacterial pathogen of major concern in the clinic) was elucidated using AP with quantitative MS [48]. Likewise, the components of close to 200 putative protein complexes were identified by AP-MS of TAP tagged proteins in the pneumonia-causing bacterial pathogen *Mycoplasma pneumoniae* [49].

Higher eukaryotic models

Global MS-based interactomic studies have also been performed in higher eukaryotic model systems. For example, AP-MS analysis of over 5,000 individual proteins that were affinity-purified from a fruit fly cell line was used to identify 556 putative protein complexes [50]. Also in this study, further experiments were performed to validate the cross-species conservation of identified PPIs by tagging close to 100 human orthologs of *Drosophila* proteins, followed by AP-MS identification of associated protein complexes in HEK293 cells. Although the test set was biased, there was an impressive 51% overlap between the original fly and the human data sets, validating the fly PPI data as a model for human inferences. Further examination of the similarity between PPIs identified in this study and publicly available interaction data reported from previous yeast and human PPI maps showed great evolutionary conservation in certain biological systems, including three major protein complexes that are involved in protein translation, protein degradation and RNA processing. In addition, p38 mitogen-activated protein kinases (MAPKs) were clearly delineated by identification of their widely interacting partners by AP-MS [51]. Analogous effective methodologies have been established for MS analysis of affinity-purified protein complexes in the multicellular nematode worm *Caenorhabditis elegans* [49,50]. The utility and evolutionary conservation of interaction networks in these and other genetically tractable metazoan organisms is well established, making them powerful models for exploring human biology and disease mechanisms [52-55].

Mouse

When considering the choice of organism for modeling human disease, the mouse is often the preferred model of choice. Yet because of the associated technical difficulties of creating large numbers of tagged mouse strains for AP-based interactomic studies, alternative approaches have to be considered for the global profiling of PPIs in mammals. Nevertheless, several recent studies have

successfully used targeted AP-based approaches followed by MS to identify select PPIs in mouse tissues or derived cell lines that are relevant to human medical conditions.

Diseases of the brain have garnered particular biomedical attention in recent years, and several mouse models of these diseases have been used in interactomic studies. For example, mouse-derived brain tissue and cell lines have been used in conjunction with AP-MS in the characterization of the interactome of LDL receptor-related protein-1 (LRP-1), a recently identified phagocytic receptor for myelin debris in the central nervous system [56]. The identified binding partners further supported the proposed role of this macrophage receptor in potentially preventing the onset of multiple sclerosis [57]. This protective role revolves around the clearance of myelin components from apoptotic oligodendrocytes, thereby preventing inflammation and an autoimmune response. Similarly, AP-MS has been used to identify proteins that are associated with huntingtin in the brain tissue of wild-type mice but not in strains carrying a mutation that causes the Huntington's disease phenotype [57]. This suggested a novel role of huntingtin in protein translation [57]. A more expansive huntingtin (htt) interactome sub-network, comprising over 700 candidate proteins, was likewise identified in mouse brain extracts using AP-MS by Shirasaki *et al.* [58]. This study did not, however, contain any experimental validation of the putative htt interactors, suggesting that the number of candidate proteins would drop following rigorous scoring and independent biological validation. Affinity purification of PSD-95 (DLG4), a membrane-bound kinase from mouse brain, allowed the identification of physically associated synaptic protein complexes that had been previously linked to schizophrenia and other diseases [59]. Likewise, the interacting partners of the prion protein, the mutant form of which forms aggregates in brain that are responsible for bovine spongiform encephalopathy (mad cow disease), were also recently tentatively identified in transgenic mice by affinity purification [60].

Other rodents represent promising models. For example, co-IP MS was applied to rat-derived myotubes to study the interactome of the insulin receptor substrate-1 protein, which plays a central role in insulin signaling and a proposed role in the development of insulin resistance in diabetes [61]. Although co-IP allowed the pull down of endogenous protein complexes directly from the tissue of interest, without the need for the genetic manipulation required for the tagging of proteins in AP-MS approaches, it must be noted that this strategy depends on the availability of a reliable antibody, whose generation, development and subsequent validation is cumbersome and time-consuming.

Mouse-derived embryonic stem (ES) and induced pluripotent stem (iPS) cells are playing an increasingly

important role as model systems for discovery studies and for screening potential therapeutics for a number of major diseases. Several interactomic studies have been performed in mES and iPS cells, complementing the molecular-profiling efforts routinely reported for these systems. The interactomes of OCT4 and SOX, two of the four 'Yamanaka' transcription factors required for the generation of pluripotent cells, were recently characterized in mouse ES cells by different AP-MS approaches [62-65]. These studies provided insight into the mechanisms of establishment and regulation of pluripotency in mouse ES cells. An analogous AP-MS study in mouse ES cells by our group, utilizing a mammalian affinity purification and lentiviral expression (MAPLE) system, was used to identify a novel link between the Klf4 reprogramming transcription factor and the chromatin-remodeling machinery that is required for efficient induction of pluripotency [66].

Human

The vast majority of MS-based studies of PPIs in human cells have been performed under tissue-culture conditions using a few representative cell lines, the vast majority of which are cancer-derived or transformed. Methodologies that can achieve high levels of coverage and recovery, similar to those provided by the large libraries of tagged proteins in yeast and *E. coli*, are being developed through the use of efficient tags and stable delivery mechanisms (such as lentivirus or clone integration) [66]. There have been several landmark studies in recent years that have contributed greatly to the mapping of a preliminary human protein interactome. Notably, Ewing *et al.* [67] selected over 300 bait proteins on the basis of their proven or predicted association with disease, transiently overexpressed them as Flag-tagged constructs in the HEK293 cell line, and then used AP-MS to identify stably associated binding partners. Following bioinformatic filtering of the initial dataset, the authors reported 6,463 high-confidence PPIs involving 2,235 human proteins. Although no biological validation experiments were performed, some of the protein complexes established in literature were identified in this study, supporting the quality of the network. Using a different co-IP-based strategy, close to 1,800 antibodies were used to identify stably interacting proteins from 3,290 immunoprecipitation pulldowns using extracts from HeLa cells, a popular cervical cancer cell line established more than 60 years ago [68].

Our own group re-analyzed both of these cell lines using an extensive chromatography-based co-fractionation strategy to enrich for stably associated protein complexes, which were subsequently identified by MS [69]. This tagless approach enabled the identification of 13,993 high-confidence physical interactions, linking

3,006 proteins as subunits of 622 putative complexes. Strikingly, the majority of the complexes, including many previously unannotated entities, had subunits that have been linked to human disease, implicating their uncharacterized binding partners as potential candidates in the same or in similar pathologies. Biochemical co-fractionation has also been used in conjunction with stable isotope labeling with amino acids in cell culture (SILAC)-based quantitative MS to study changes in the abundance of soluble cytosolic protein complexes in HeLa cells in response to growth-factor treatment [70].

In addition to the global interactome studies outlined above, there have been several targeted studies examining particular protein associations in specific diseases. For example, TAP-analysis of SCRIB, a protein important in the development of cell polarity, was used to identify a protein complex that is associated with the metastatic progression of breast cancer [71]. AP-MS was also used to isolate and identify proteins that are associated with tagged versions of lebercilin, with the aim of determining the functional consequences of mutations in this protein, which are responsible for the development of Leber congenital amaurosis (a disease causing childhood blindness) [72]. The study provided insights into the molecular mechanisms associated with normal ciliary function and into perturbations that are linked to disease. Co-IP MS identification of proteins from cardiac and skeletal muscle that interact with dystrophin (a protein responsible for a number of myopathies) has also led to the identification of tissue-specific signaling pathways that seem to play a role in cardiac disease and muscular dystrophy [73].

By and large, most of the PPIs reported to date have been studied experimentally in human cancer cell lines. For example, functionally relevant interactors of a mutant p53 protein variant previously shown to increase tumor invasion and metastasis in mice were identified by co-IP-MS in cancer cell lines [74]. Likewise, affinity purification of tagged EGFR (a cell surface receptor that is overexpressed in a number of cancers) led to the identification and quantification (by isobaric tags for relative and absolute quantification (iTRAQ)-based stable isotope labeling) of differential binding partners in lung tumor cell lines [75]. Several proteins with potentially crucial roles in the development of melanoma were elucidated by AP-MS analysis of hypoxia induced factor 2 (HIF2, a transcription factor commonly overexpressed in aggressive cancers) in human melanoma cell lines [76]. Likewise, novel interactors of the adenomatous polyposis coli (APC) oncoprotein were identified by AP-MS in HEK293 cells [77]. Collectively, these studies provided new candidate co-factors of regulators of systems commonly disrupted in cancer.

AP-MS analysis of human cell line models has also been used to monitor the impact of drug treatment on

PPI networks and protein complexes. For example, the interactome of the estrogen receptor alpha (ER alpha), a crucial transcription factor in hormone-responsive breast cancer, was analyzed by AP-MS after treating breast cancer cells with three different therapeutic antagonistic ligands in comparison to an agonist [78]. This led to the identification of novel nuclear cofactors for ER alpha, each of which was active when the receptor was bound to a different estrogen antagonist, providing further understanding of their different pharmacological properties. The interactomes of the p53/p63 master tumor suppressor regulators were also recently mapped by AP-MS in cisplatin-treated squamous cell carcinoma cells, thereby probing their involvement in the development of resistance to this chemotherapy [79]. A combination of AP and quantitative MS was also used recently to examine the target-binding specificity of 16 different histone deacetylase (HDAC) inhibitors that have therapeutic potential as anti-cancer drugs [80], with the differences in observed binding profiles supporting unique modes of action.

Bioinformatics from global proteomic and genomic data

Given the difficulties associated with scaling up interaction experiments, the analysis of PPI networks using bioinformatic methods is increasingly popular. One of the most commonly used tools for the visualization and integration of PPI networks is Cytoscape. There are close to 160 publicly available plugins for additional data analysis within this open-source software suite [81]. In general, the source data used in computational approaches to evaluate PPI and even to predict interaction maps come from global mRNA expression-profiling studies. These rely on information from curated interaction databases, populated to a large degree by experimental data emerging from two-hybrid studies, both for scoring and benchmarking the PPI predictions. There are several publicly available databases that contain predictive and experimental PPI information, including Biological General Repository for Interaction Datasets (BioGRID), Database of Interacting Proteins (DIP), Molecular Interactions Database (MINT), and Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) [1,12]. Other available PPI databases and methods of predicting PPI networks have been reviewed recently by Liu and Chen [82].

Nevertheless, experimentally confirmed PPIs stemming from two-hybrid system, AP-MS and small-scale interaction studies account for less than 25% of all human PPIs predicted by certain sources [83]. This gap in knowledge has motivated the development of innovative computational procedures for *de novo* prediction of PPIs, which are not based on direct experimental evidence.

Computational methods can utilize existing genomic knowledge of gene and protein evolutionary conservation, gene neighborhoods, subcellular localization, co-expression, structural similarity and docking compatibility to predict PPI networks. The prediction of PPI networks on the basis of AP-MS and other high-throughput data have been reviewed recently [82,84]. Several recent studies have showcased the scope for computational modeling. One modeled a network containing over 94,000 PPIs (462 of which were verified by independent yeast two-hybrid and quantitative MS-based experiments) that implicated TOMM40 as a potential factor in Alzheimer's disease [85,86]. Another identified novel PPIs driving apoptosis by prediction based on three-dimensional structures of protein complexes in this pathway [87]. Therefore, there is high hope that closer integration of computational methods and experimental validation can be used to produce reliable PPI networks that will provide a more extensive picture of differences between 'normal' and disease-perturbed proteomes.

Global predictive studies of these types have been used with some success in delineating potentially clinically informative interactions. For example, disease progression in and clinical outcomes of breast cancer were predicted in a pioneering study based on examining changes in the connectivity of 'hub' proteins in tumor cells. Existing PPI literature and curated databases were searched and the networks within them overlaid on the public gene expression data to define two different types of PPI modules, those that have protein interactors that are co-expressed only in a specific tissue and those that are co-expressed in all or most tissues [88]. Using gene expression data from breast adenocarcinoma patient samples, changes in these modules were found to be highly predictive of cancer progression and patient morbidity. In another analogous recent study, existing PPI information from databases and gene expression data from patients with aggressive and indolent chronic lymphocytic leukemia were used to predict 38 PPI subnetworks indicative of disease progression [89]. Integrative bioinformatic analysis of gene expression data with existing PPI information has also been used to show that human tissue developmental processes, breast cancer prognosis and the progression of brain cancers reflect a compendium of competing interactions resulting from the combined actions of differentially expressed protein subnetworks [90].

Conclusions

Studies of PPI networks and protein complexes have been performed, to varying extents, on all levels of life, from viruses and unicellular organisms to mammalian model systems and human tissues. To gain the maximal amount of biomedically relevant information, each of these

studies should not be looked at separately because information useful to clinical applications can potentially be found in each model system. The scope of the yeast and bacterial AP-MS datasets and the experimental versatility of these organisms, in terms of genetic manipulation and established methodologies and resource databases, have proven to be indispensable in the development of the basic technologies and bioinformatic approaches used in the study of physical interaction networks and in identifying PPI that are conserved at all levels of life. This has led to a number of analogous interactomic approaches in higher level eukaryotes, allowing for a better understanding of the composition of stable protein complexes and their functional relevance in human disease contexts. The lessons learned from these model systems have begun to be applied in the analysis of human disease networks, with the ultimate goal of porting the analysis directly to clinical samples.

It must be noted that AP-MS approaches often suffer from several significant limitations stemming from the fact that samples produced by affinity purifications contain not only interacting proteins but also proteins that are non-specifically bound to the affinity matrix and other common contaminants resulting from limitations in the enrichment procedure. This results in potentially high false-positive rates. The solution to this issue can partly be found in stringent washing of non-specific binders, but at the cost of losing weak interactions. Dual-step TAP methods can also alleviate this issue but often require large amounts of sample because of losses at each stage. Therefore, stringent controls for the purpose of identifying non-specific binders, computational filtering and independent PPI validation methods are required. The gold standard for the validation of interactions is the IP-western, but with the recent advances in quantitative targeted proteomics in addition to the discovery of PPIs, MS-based methods can now be used for validation studies. Recent applications of the SRM and SWATH methodologies for the discovery and confirmation of interactions with the Grb2 signaling protein can serve as prime examples of strategies for dealing with this complexity of cell systems [91,92].

There are several other major challenges that must be tackled in the coming years, most technical but some computational. These include the need for more comprehensive experimental mapping of lower-abundance protein assemblies and transient PPIs for the purpose of creating more extensive databases of verified PPIs, the development of novel high-throughput, reliable PPI mapping methodologies that could be applied to clinically relevant samples directly, and improvements in bioinformatic analysis and data integration from multiple sources. These three streams of research are proceeding hand-in-hand in our laboratory and many others, and are

reliant to a large degree on the model systems being used, each with their inherent advantages and limitations. The next great step in the field will be a move to engage and inspire clinicians to see the value of measuring interaction networks under normal and disease states, as well as the targeting of PPIs by therapeutics and the monitoring of PPI patterns as potential outputs in diagnostic and prognostic screens. Given that the initial steps towards these goals are well under way, the active promotion of translational biomedical problems in research institutions across the world will only help the cause.

Competing interests

The authors declare that they have no competing interests.

Abbreviations

AP, affinity purification; co-IP, co-immunoprecipitation; ES, embryonic stem; htt, huntingtin; iPS, induced pluripotent stem; MS, mass spectrometry; PPI, protein-protein interaction; SRM, selected reaction monitoring; TAP, tandem affinity purification.

Published: 30 April 2013

References

1. Bonetta L: **Protein-protein interactions: interactome under construction.** *Nature* 2010, **468**:851-854.
2. Venkatesan K, Rual JF, Vazquez A, Stelzl U, Lemmens I, Hirozane-Kishikawa T, Hao T, Zenkner M, Xin X, Goh KI, Yildirim MA, Simonis N, Heinzmann K, Gebreab F, Sahalie JM, Cevik S, Simon C, de Smet AS, Dann E, Smolyar A, Vinayagam A, Yu H, Szeto D, Borick H, Dricot A, Klitgord N, Murray RR, Lin C, Lalowski M, Timm J, *et al.*: **An empirical framework for binary interactome mapping.** *Nat Methods* 2009, **6**:83-90.
3. Brown CJ, Lain S, Verma CS, Fersht AR, Lane DP: **Awakening guardian angels: drugging the p53 pathway.** *Nat Rev Cancer* 2009, **9**:862-873.
4. Joerger AC, Fersht AR: **Structure-function-rescue: the diverse nature of common p53 cancer mutants.** *Oncogene* 2007, **26**:2226-2242.
5. Brooke MA, Nitoiu D, Kelsell DP: **Cell-cell connectivity: desmosomes and disease.** *J Pathol* 2012, **226**:158-171.
6. Forman MS, Trojanowski JQ, Lee VM: **Neurodegenerative diseases: a decade of discoveries paves the way for therapeutic breakthroughs.** *Nat Med* 2004, **10**:1055-1063.
7. Ross CA, Poirier MA: **Protein aggregation and neurodegenerative disease.** *Nat Med* 2004, **Suppl 10**:S10-17.
8. Ma-Lauer Y, Lei J, Hilgenfeld R, von Brunn A: **Virus-host interactomes - antiviral drug discovery.** *Curr Opin Virol* 2012, **2**:614-621.
9. Zheng J, Sugrue RJ, Tang K: **Mass spectrometry based proteomic studies on viruses and hosts - a review.** *Anal Chim Acta* 2011, **702**:149-159.
10. Kaake RM, Wang X, Huang L: **Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry.** *Mol Cell Proteomics* 2010, **9**:1650-1665.
11. Ramisetty SR, Washburn MP: **Unraveling the dynamics of protein interactions with quantitative mass spectrometry.** *Crit Rev Biochem Mol Biol* 2011, **46**:216-228.
12. Sardi ME, Washburn MP: **Building protein-protein interaction networks with proteomics and informatics tools.** *J Biol Chem* 2011, **286**:23645-23651.
13. Gavin AC, Maeda K, Kuhner S: **Recent advances in charting protein-protein interaction: mass spectrometry-based approaches.** *Curr Opin Biotechnol* 2011, **22**:42-49.
14. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzlaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Toksoz E, Droegge A, Krobitsch S, Korn B, Birchmeier W, Lehrach H, Wanker EE: **A human protein-protein interaction network: a resource for annotating the proteome.** *Cell* 2005, **122**:957-968.
15. Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S,

- Albala JS, Lim J, Fraughton C, Llamasos E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, et al: **Towards a proteome-scale map of the human protein-protein interaction network.** *Nature* 2005, **437**:1173-1178.
16. Stynen B, Tournu H, Tavernier J, Van Dijk P: **Diversity in genetic *in vivo* methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system.** *Microbiol Mol Biol Rev* 2012, **76**:331-382.
17. Gingras AC, Gstaiger M, Raught B, Aebersold R: **Analysis of protein complexes using mass spectrometry.** *Nat Rev Mol Cell Biol* 2007, **8**:645-654.
18. Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R: **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* 2012, **11**:O111 016717.
19. Picotti P, Aebersold R: **Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions.** *Nat Methods* 2012, **9**:555-566.
20. Bantscheff M, Lemeer S, Savitski MM, Kuster B: **Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present.** *Anal Bioanal Chem* 2012, **404**:939-965.
21. Babu M, Vlasblom J, Pu S, Guo X, Graham C, Bean BD, Burston HE, Vizeacoumar FJ, Snider J, Phanse S, Fong V, Tam YY, Davey M, Hnatshak O, Bajaj N, Chandran S, Punna T, Christopoulos C, Wong V, Yu A, Zhong G, Li J, Stagljar I, Conibear E, Wodak SJ, Emili A, Greenblatt JF: **Interaction landscape of membrane-protein complexes in *Saccharomyces cerevisiae*.** *Nature* 2012, **489**:585-589.
22. Volkel P, Le Faou P, Angrand PO: **Interaction proteomics: characterization of protein complexes using tandem affinity purification-mass spectrometry.** *Biochem Soc Trans* 2010, **38**:883-887.
23. Xu X, Song Y, Li Y, Chang J, Zhang H, An L: **The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification.** *Protein Expr Purif* 2010, **72**:149-156.
24. Babu M, Butland G, Pogoutse O, Li J, Greenblatt JF, Emili A: **Sequential peptide affinity purification system for the systematic isolation and identification of protein complexes from *Escherichia coli*.** *Methods Mol Biol* 2009, **564**:373-400.
25. Zeghouf M, Li J, Butland G, Borkowska A, Canadien V, Richards D, Beattie B, Emili A, Greenblatt JF: **Sequential Peptide Affinity (SPA) system for the identification of mammalian and bacterial protein complexes.** *J Proteome Res* 2004, **3**:463-468.
26. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, et al: **Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*.** *Nature* 2006, **440**:637-643.
27. Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J, Emili A: **Interaction network containing conserved and essential protein complexes in *Escherichia coli*.** *Nature* 2005, **433**:531-537.
28. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, et al: **Functional organization of the yeast proteome by systematic analysis of protein complexes.** *Nature* 2002, **415**:141-147.
29. Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutillier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreau M, Muskaf B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, et al: **Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry.** *Nature* 2002, **415**:180-183.
30. Ito T, Tashiro K, Muta S, Ozawa R, Chiba T, Nishizawa M, Yamamoto K, Kuhara S, Sakaki Y: **Toward a protein-protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins.** *Proc Natl Acad Sci U S A* 2000, **97**:1143-1147.
31. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B: **A generic protein purification method for protein complex characterization and proteome exploration.** *Nat Biotechnol* 1999, **17**:1030-1032.
32. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamar G, Yang M, Johnston M, Fields S, Rothberg JM: **A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*.** *Nature* 2000, **403**:623-627.
33. von Mering C, Krause R, Snel B, Cornell M, Oliver SG, Fields S, Bork P: **Comparative assessment of large-scale data sets of protein-protein interactions.** *Nature* 2002, **417**:399-403.
34. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dimpfelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, et al: **Proteome survey reveals modularity of the yeast cell machinery.** *Nature* 2006, **440**:631-636.
35. Breitkreutz A, Choi H, Sharom JR, Boucher L, Neduva V, Larsen B, Lin ZY, Breitkreutz BJ, Stark C, Liu G, Ahn J, Dewar-Darch D, Reguly T, Tang X, Almeida R, Qin ZS, Pawson T, Gingras AC, Nesvizhskii AI, Tyers M: **A global protein kinase and phosphatase interaction network in yeast.** *Science* 2010, **328**:1043-1046.
36. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K, Hernandez H, Jang GM, Roth SL, Akiva E, Marlett J, Stephens M, D'Orso I, Fernandes J, Fahey M, Mahon C, O'Donoghue AJ, Todorovic A, Morris JH, Maltby DA, Alber T, Cagney G, Bushman FD, Young JA, Chanda SK, Sundquist WL, et al: **Global landscape of HIV-human protein complexes.** *Nature* 2011, **481**:365-370.
37. Jager S, Gulbahce N, Cimermancic P, Kane J, He N, Chou S, D'Orso I, Fernandes J, Jang G, Frankel AD, Alber T, Zhou Q, Krogan NJ: **Purification and characterization of HIV-human protein complexes.** *Methods* 2011, **53**:13-19.
38. Raghavendra NK, Shkriabai N, Graham R, Hess S, Kvaratskhelia M, Wu L: **Identification of host proteins associated with HIV-1 preintegration complexes isolated from infected CD4+ cells.** *Retrovirology* 2010, **7**:66.
39. Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A, Burckstummer T, Muller AC, Fauster A, Holze C, Lindsten K, Goodbourn S, Kochs G, Weber F, Bartschschlager R, Bowie AG, Bennett KL, Colinge J, Superti-Furga G: **Viral immune modulators perturb the human molecular network by common and unique strategies.** *Nature* 2012, **487**:486-490.
40. Komarova AV, Combredet C, Meyniel-Schicklin L, Chapelle M, Caignard G, Camadro JM, Lotteau V, Vidalain PO, Tangy F: **Proteomic analysis of virus-host interactions in an infectious context using recombinant viruses.** *Mol Cell Proteomics* 2011, **10**:M110 007443.
41. Chatel-Chaix L, Melancon P, Racine ME, Baril M, Lamarre D: **Y-box-binding protein 1 interacts with hepatitis C virus NS3/4A and influences the equilibrium between viral RNA replication and infectious particle production.** *J Virol* 2011, **85**:11022-11037.
42. Katoh H, Okamoto T, Fukuhara T, Kambara H, Morita E, Mori Y, Kamitani W, Matsuura Y: **Japanese encephalitis virus core protein inhibits stress granule formation through an interaction with Caprin-1 and facilitates viral propagation.** *J Virol* 2013, **87**:489-502.
43. Hu P, Janga SC, Babu M, Diaz-Mejia JJ, Butland G, Yang W, Pogoutse O, Guo X, Phanse S, Wong P, Chandran S, Christopoulos C, Nazarians-Armavil A, Nasserri NK, Musso G, Ali M, Nazemof N, Eroukova V, Golshani A, Paccanaro A, Greenblatt JF, Moreno-Hagelsieb G, Emili A: **Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins.** *PLoS Biol* 2009, **7**:e96.
44. Maddalo G, Stenberg-Bruzell F, Gotzke H, Toddo S, Bjorkholm P, Eriksson H, Chovanec P, Genevaux P, Lehtio J, Ilag LL, Daley DO: **Systematic analysis of native membrane protein complexes in *Escherichia coli*.** *J Proteome Res* 2011, **10**:1848-1859.
45. Yang X, Promnarek K, Qin J, He M, Shroder DY, Kariu T, Wang Y, Pal U: **Characterization of multiprotein complexes of the *Borrelia burgdorferi* outer membrane vesicles.** *J Proteome Res* 2011, **10**:4556-4566.
46. Liu X, Afrane M, Clemmer DE, Zhong G, Nelson DE: **Identification of *Chlamydia trachomatis* outer membrane complex proteins by differential proteomics.** *J Bacteriol* 2010, **192**:2852-2860.
47. Marzoo J, Sanchez S, Ferreiros CM, Criado MT: **Identification of *Neisseria meningitidis* outer membrane vesicle complexes using 2-D high resolution clear native/SDS-PAGE.** *J Proteome Res* 2010, **9**:611-619.
48. Cherkasov A, Hsing M, Zoraghi R, Foster LJ, See RH, Stoynov N, Jiang J, Kaur S, Lian T, Jackson L, Gong H, Swayze R, Amandoron E, Hormozdiari F, Dao P, Sahinalp C, Santos-Filho O, Axerio-Cilies P, Byler K, McMaster WR, Brunham RC, Finlay BB, Reiner NE: **Mapping the protein interaction network in**

- methicillin-resistant *Staphylococcus aureus*. *J Proteome Res* 2011, **10**:1139-1150.
49. Kuhner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, Rode M, Yamada T, Maier T, Bader S, Beltran-Alvarez P, Castano-Diez D, Chen WH, Devos D, Guell M, Norambuena T, Racke I, Rybin V, Schmidt A, Yus E, Aebersold R, Herrmann R, Bottcher B, Frangakis AS, Russell RB, Serrano L, Bork P, Gavin AC: **Proteome organization in a genome-reduced bacterium**. *Science* 2009, **326**:1235-1240.
 50. Guruharsha KG, Rual JF, Zhai B, Mintseris J, Vaidya P, Vaidya N, Beekman C, Wong C, Rhee DY, Cenaj O, McKillip E, Shah S, Stapleton M, Wan KH, Yu C, Parsa B, Carlson JW, Chen X, Kapadia B, VijayRaghavan K, Gygi SP, Celniker SE, Obar RA, Artavanis-Tsakonas S: **A protein complex network of *Drosophila melanogaster***. *Cell* 2011, **147**:690-703.
 51. Belozero V, Lin ZY, Gingras AC, McDermott JC, Michael Siu KW: **High-resolution protein interaction map of the *Drosophila melanogaster* p38 mitogen-activated protein kinases reveals limited functional redundancy**. *Mol Cell Biol* 2012, **32**:3695-3706.
 52. Aboobaker AA, Blaxter ML: **Medical significance of *Caenorhabditis elegans***. *Ann Med* 2000, **32**:23-30.
 53. Fortini ME, Skupski MP, Boguski MS, Hariharan IK: **A survey of human disease gene counterparts in the *Drosophila* genome**. *J Cell Biol* 2000, **150**:F23-F30.
 54. Lehner B: **Genotype to phenotype: lessons from model organisms for human genetics**. *Nat Rev Genet* 2013, **14**:168-178.
 55. Pradel E, Ewbank JJ: **Genetic models in pathogenesis**. *Annu Rev Genet* 2004, **38**:347-363.
 56. Fernandez-Castaneda A, Arandjelovic S, Stiles TL, Schlobach RK, Mowen KA, Gonias SL, Gaultier A: **Identification of the LDL receptor-related protein-1 interactome in central nervous system myelin suggests a role in the clearance of necrotic cell debris**. *J Biol Chem* 2013, **288**:4538-4548.
 57. Culver BP, Savas JN, Park SK, Choi JH, Zheng S, Zeitlin SO, Yates JR 3rd, Tanese N: **Proteomic analysis of wild-type and mutant huntingtin-associated proteins in mouse brains identifies unique interactions and involvement in protein synthesis**. *J Biol Chem* 2012, **287**:21599-21614.
 58. Shirasaki DI, Greiner ER, Al-Ramahi I, Gray M, Boontheung P, Geschwind DH, Botas J, Coppola G, Horvath S, Loo JA, Yang XW: **Network organization of the huntingtin proteomic interactome in mammalian brain**. *Neuron* 2012, **75**:41-57.
 59. Fernandez E, Collins MO, Uren RT, Kopanitsa MV, Komiyama NH, Croning MD, Zografos L, Armstrong JD, Choudhary JS, Grant SG: **Targeted tandem affinity purification of PSD-95 recovers core postsynaptic complexes and schizophrenia susceptibility proteins**. *Mol Syst Biol* 2009, **5**:269.
 60. Rutishauser D, Mertz KD, Moos R, Brunner E, Rulicke T, Calella AM, Aguzzi A: **The comprehensive native interactome of a fully functional tagged prion protein**. *PLoS One* 2009, **4**:e4446.
 61. Geetha T, Langlais P, Luo M, Mapes R, Lefort N, Chen SC, Mandarino LJ, Yi Z: **Label-free proteomic identification of endogenous, insulin-stimulated interaction partners of insulin receptor substrate-1**. *J Am Soc Mass Spectrom* 2011, **22**:457-466.
 62. Fang X, Yoon JG, Li L, Tsai YS, Zheng S, Hood L, Goodlett DR, Foltz G, Lin B: **Landscape of the SOX2 protein-protein interactome**. *Proteomics* 2011, **11**:921-934.
 63. Gao Z, Cox JL, Gilmore JM, Ormsbee BD, Mallanna SK, Washburn MP, Rizzino A: **Determination of protein interactome of transcription factor Sox2 in embryonic stem cells engineered for inducible expression of four reprogramming factors**. *J Biol Chem* 2012, **287**:11384-11397.
 64. Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary J: **An expanded Oct4 interaction network: implications for stem cell biology, development, and disease**. *Cell Stem Cell* 2010, **6**:382-395.
 65. van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA: **An Oct4-centered protein interaction network in embryonic stem cells**. *Cell Stem Cell* 2010, **6**:369-381.
 66. Mak AB, Ni Z, Hewel JA, Chen GI, Zhong G, Karamboulas K, Blakely K, Smiley S, Marcon E, Roudeva D, Li J, Olsen JB, Wan C, Punna T, Isserlin R, Chetyrkin S, Gingras AC, Emili A, Greenblatt J, Moffat J: **A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency**. *Mol Cell Proteomics*, **9**:811-823.
 67. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M, Taylor R, Dharsee M, Ho Y, Heilbut A, Moore L, Zhang S, Ornatsky O, Bukhman YV, Ethier M, Sheng Y, Vasilescu J, Abu-Farha M, Lambert JP, Duestel HS, Stewart I, Kuehl B, Hogue K, Colwill K, Gladwish K, Muskat B, et al.: **Large-scale mapping of human protein-protein interactions by mass spectrometry**. *Mol Syst Biol* 2007, **3**:89.
 68. Malovannaya A, Lanz RB, Jung SY, Bulynko Y, Le NT, Chan DW, Ding C, Shi Y, Yucer N, Krenciute G, Kim BJ, Li C, Chen R, Li W, Wang Y, O'Malley BW, Qin J: **Analysis of the human endogenous coregulator complexome**. *Cell* 2011, **145**:787-799.
 69. Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL, Li Z, Wang PI, Boutz DR, Fong V, Phanse S, Babu M, Craig SA, Hu P, Wan C, Vlasblom J, Dar VU, Bezginov A, Clark GW, Wu GC, Wodak SJ, Tillier ER, Paccanaro A, Marcotte EM, Emili A: **A census of human soluble protein complexes**. *Cell* 2012, **150**:1068-1081.
 70. Kristensen AR, Gsponer J, Foster LJ: **A high-throughput approach for measuring temporal changes in the interactome**. *Nat Methods* 2012, **9**:907-909.
 71. Anastas JN, Biechele TL, Robitaille M, Muster J, Allison KH, Angers S, Moon RT: **A protein complex of SCRIB, NOS1AP and VANGL1 regulates cell polarity and migration, and is associated with breast cancer progression**. *Oncogene* 2012, **31**:3696-3708.
 72. Boldt K, van Reeuwijk J, Gloeckner CJ, Ueffing M, Roepman R: **Tandem affinity purification of ciliopathy-associated protein complexes**. *Methods Cell Biol* 2009, **91**:143-160.
 73. Johnson EK, Zhang L, Adams ME, Phillips A, Freitas MA, Froehner SC, Green-Church KB, Montanaro F: **Proteomic analysis reveals new cardiac-specific dystrophin-associated proteins**. *PLoS One* 2012, **7**:e43515.
 74. Coffill CR, Muller PA, Oh HK, Neo SP, Hogue KA, Cheek CF, Vousden KH, Lane DP, Blackstock WP, Gunaratne J: **Mutant p53 interactome identifies nardilysin as a p53R273H-specific binding partner that promotes invasion**. *EMBO Rep* 2012, **13**:638-644.
 75. Haura EB, Muller A, Breitwieser FP, Li J, Grebien F, Colinge J, Bennett KL: **Using iTRAQ combined with tandem affinity purification to enhance low-abundance proteins associated with somatically mutated EGFR core complexes in lung cancer**. *J Proteome Res* 2011, **10**:182-190.
 76. Steunou AL, Ducoux-Petit M, Lazar I, Monsarrat B, Erard M, Muller C, Clottes E, Burlet-Schiltz O, Nieto L: **Identification of HIF2 alpha nuclear interactome in melanoma cells reveals master proteins involved in melanoma development**. *Mol Cell Proteomics* 2013, **12**:736-748.
 77. Song J, Hao Y, Du Z, Wang Z, Ewing RM: **Identifying novel protein complexes in cancer cells using epitope-tagging of endogenous human genes and affinity-purification mass spectrometry**. *J Proteome Res* 2012, **11**:5630-5641.
 78. Cirillo F, Nassa G, Tarallo R, Stellato C, Filippo MR, Ambrosino C, Baumann M, Nyman TA, Weisz A: **Molecular mechanisms of selective estrogen receptor modulator activity in human breast cancer cells: identification of novel nuclear cofactors of antiestrogen-ER alpha complexes by interaction proteomics**. *J Proteome Res* 2013, **12**:421-431.
 79. Huang Y, Jeong JS, Okamura J, Sook-Kim M, Zhu H, Guerrero-Preston R, Ratovitski EA: **Global tumor protein p53/p63 interactome: making a case for cisplatin chemoresistance**. *Cell Cycle* 2012, **11**:2367-2379.
 80. Bantscheff M, Hopf C, Savitski MM, Dittmann A, Grandi P, Michon AM, Schlegl J, Abraham Y, Becher I, Bergamini G, Boesche M, Dellling M, Dumpelfeld B, Eberhard D, Huthmacher C, Mathieson T, Poelckel D, Reader V, Strunk K, Sweetman G, Kruse U, Neubauer G, Ramsden NG, Drewes G: **Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes**. *Nat Biotechnol* 2011, **29**:255-265.
 81. Saito R, Smoot ME, Ono K, Ruschekinski J, Wang PL, Lotia S, Pico AR, Bader GD, Ideker T: **A travel guide to Cytoscape plugins**. *Nat Methods* 2012, **9**:1069-1076.
 82. Liu ZP, Chen L: **Proteome-wide prediction of protein-protein interactions from high-throughput data**. *Protein Cell* 2012, **3**:508-520.
 83. Elefsinioti A, Sarac OS, Hegele A, Plake C, Hubner NC, Poser I, Sarov M, Hyman A, Mann M, Schroeder M, Stelzl U, Beyer A: **Large-scale de novo prediction of physical protein-protein association**. *Mol Cell Proteomics* 2011, **10**:M111 010629.
 84. Choi H: **Computational detection of protein complexes in AP-MS experiments**. *Proteomics* 2012, **12**:1663-1668.
 85. Kuzu G, Keskin O, Gursoy A, Nussinov R: **Constructing structural networks of signaling pathways on the proteome scale**. *Curr Opin Struct Biol* 2012, **22**:367-377.
 86. Tuncbag N, Gursoy A, Nussinov R, Keskin O: **Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM**. *Nat Protoc* 2011, **6**:1341-1354.
 87. Acuner Ozbabacan SE, Keskin O, Nussinov R, Gursoy A: **Enriching the human apoptosis pathway by predicting the structures of protein-protein**

- complexes. *J Struct Biol* 2012, **179**:338-346.
88. Taylor IW, Linding R, Warde-Farley D, Liu Y, Pesquita C, Faria D, Bull S, Pawson T, Morris Q, Wrana JL: **Dynamic modularity in protein interaction networks predicts breast cancer outcome.** *Nat Biotechnol* 2009, **27**:199-204.
89. Chuang HY, Rassenti L, Salcedo M, Licon K, Kohlmann A, Haferlach T, Foa R, Ideker T, Kipps TJ: **Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression.** *Blood* 2012, **120**:2639-2649.
90. Dutkowski J, Ideker T: **Protein networks as logic functions in development and cancer.** *PLoS Comput Biol* 2011, **7**:e1002180.
91. Bisson N, James DA, Ivosev G, Tate SA, Bonner R, Taylor L, Pawson T: **Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor.** *Nat Biotechnol* 2011, **29**:653-658.
92. Findlay GM, Smith MJ, Lanner F, Hsiung MS, Gish GD, Petsalaki E, Cockburn K, Kaneko T, Huang H, Bagshaw RD, Ketela T, Tucholska M, Taylor L, Bowtell DD, Moffat J, Ikura M, Li SS, Sidhu SS, Rossant J, Pawson T: **Interaction domains of Sos1/Grb2 are finely tuned for cooperative control of embryonic stem cell fate.** *Cell* 2013, **152**:1008-1020.

doi:10.1186/gm441

Cite this article as: Kuzmanov U, Emili A: **Protein-protein interaction networks: probing disease mechanisms using model systems.** *Genome Medicine* 2013, **5**:37.