Cell
Leading Edge

Review

Discovery and significance of protein-protein interactions in health and disease

Jack F. Greenblatt,[1,2,*] Bruce M. Alberts,[3] and Nevan J. Krogan[4,5,6,*]
[1]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[2]Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada
[3]Department of Biochemistry and Biophysics, University of California, San Francisco (UCSF), San Francisco, CA, USA
[4]Quantitative Biosciences Institute, University of California, San Francisco (UCSF), San Francisco, CA, USA
[5]Gladstone Institute of Data Science and Biotechnology, J. David Gladstone Institutes, San Francisco, CA, USA
[6]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco (UCSF), San Francisco, CA, USA
*Correspondence: jack.greenblatt@utoronto.ca (J.F.G.), nevan.krogan@ucsf.edu (N.J.K.)
https://doi.org/10.1016/j.cell.2024.10.038

SUMMARY

The identification of individual protein-protein interactions (PPIs) began more than 40 years ago, using protein affinity chromatography and antibody co-immunoprecipitation. As new technologies emerged, analysis of PPIs increased to a genome-wide scale with the introduction of intracellular tagging methods, affinity purification (AP) followed by mass spectrometry (MS), and co-fractionation MS (CF-MS). Now, combining the resulting catalogs of interactions with complementary methods, including crosslinking MS (XL-MS) and cryogenic electron microscopy (cryo-EM), helps distinguish direct interactions from indirect ones within the same or between different protein complexes. These powerful approaches and the promise of artificial intelligence applications like AlphaFold herald a future where PPIs and protein complexes, including energy-driven protein machines, will be understood in exquisite detail, unlocking new insights in the contexts of both basic biology and disease.

INTRODUCTION

The sophisticated chemistry that makes life possible is determined by the biological macromolecules that constitute nearly 90% of the dry mass of a cell. Predominant among these macromolecules are proteins, the remarkably versatile polymers that both catalyze most of the many chemical reactions needed in the cell and play critical structural roles. Both types of functions often require that proteins assemble with each other to form protein complexes. Many of these function as "protein machines" in which ordered, allosteric shape changes produce a series of conformational changes useful for the cell.[1] Some protein machines are understood in great detail, having been reconstituted from pure preparations of their components and investigated extensively *in vitro*, eventually leading to high-resolution structural characterization. Two well-known examples are the large multi-protein complexes that carry out DNA transcription[2,3] and DNA replication.[4] In these cases, the discovery and dissection of the protein machines required that sensitive and highly specific assays for particular biochemical processes be created. Combined with genetic dissections, this then allowed all of the components that catalyze the process inside of the cell to be identified, purified, and functionally characterized in a test tube. Similar biochemical assays for function allowed researchers to develop a molecular and structural understanding of numerous other protein machines, including the proteasome,[5,6] chromatin remodeling complexes,[7] the spliceosome,[8,9] and the ribosome[10–13] (these last two being machines in which RNAs play a central role).

It is sobering to reflect on the fact that the simplest known living cell, an engineered mycoplasma, contains only 473 genes that are important for its reproduction, and yet nearly a third have no known function.[14] For the human genome, even ignoring the problem that many genes have multiple functions, we can only guess about the function for many thousands of our ~20,000 protein-coding genes.[15,16] How many activities are needed for life to exist about which we have not the slightest clue? Hundreds of other, as yet uncharacterized, protein machines likely exist in our cells for which we have no biochemical assay, with a different approach being needed to find them. Understanding them will be critical for our understanding of both fundamental biological processes and the biology behind disease states—knowledge that is crucial for the development of badly needed therapeutics. Protein-protein interaction (PPI) maps represent a promising way of attacking these problems.

INSIGHTS INTO BIOLOGICAL FUNCTION FROM THE IDENTIFICATION OF PPIs USING PROTEIN AFFINITY CHROMATOGRAPHY

Advances in understanding cellular PPIs have gone lock step with technological innovations (Figure 1). Because the interaction of proteins with other molecules often has exquisite specificity, some early attempts to identify PPIs utilized protein affinity chromatography. Affinity chromatography has been used since

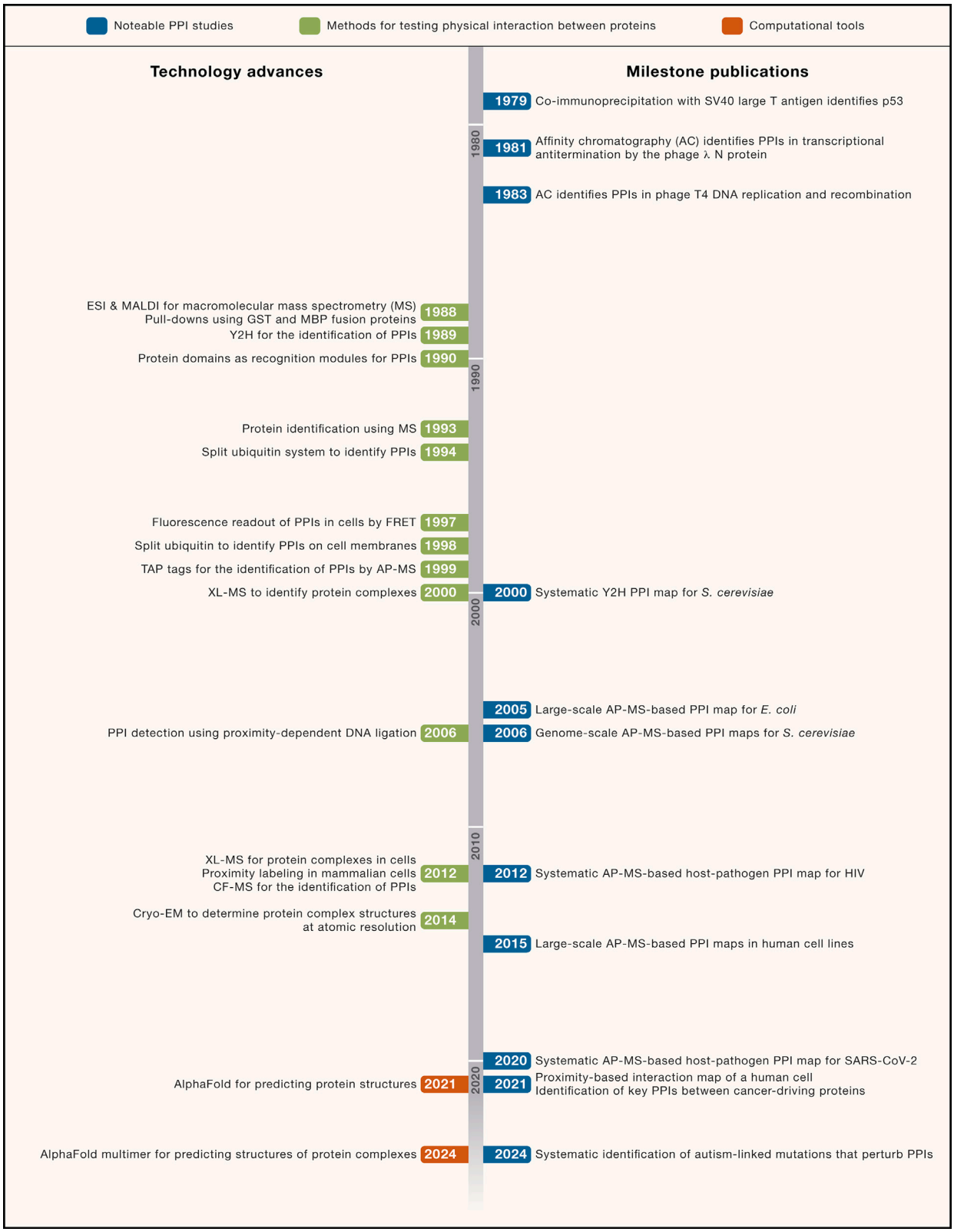

Figure 1. Timeline of notable studies and experimental and computational methods for assessing protein-protein interactions

the early twentieth century as a method to purify proteins that bind to a particular molecule or to characterize interactions of enzymes and other proteins with their substrates, cofactors, and other interactors.[17,18] It was probably first used in 1910 by Starkenstein to purify α-amylase by using insoluble starch as both the ligand and support.[19] In the ensuing decades, many improvements were made in the supports that were used for affinity chromatography and in the methods that were used to couple ligands to them,[17,18] often with the objective of using immobilized antigens to purify specific antibodies from serum. At first, antigens were coupled to supports like kaolin or charcoal,[20,21] but an important improvement was the use of diazotized *p*-aminobenzyl-cellulose for the covalent coupling first of bovine serum albumin and then of various haptens for the isolation of specific antibodies.[22,23] Still later, the introduction of beaded agarose (Sepharose), rather than cellulose, improved the mechanical properties of the support,[24] and the introduction of cyanogen bromide (CNBr) activation of the agarose beads made it easy to couple proteins and other molecules to agarose via primary amino groups.[25] Enzyme substrates[26] and protein antigens coupled to CNBr-activated agarose were then used routinely to purify enzymes and antibodies.

Although, in principle, a purified protein immobilized in this way could be used to search in an unbiased way for its unknown interaction partners in cell extracts, that was not done until the early 1980s, when protein affinity chromatography[27] was used to search for proteins that interacted with bacteriophage λ and bacteriophage T4 proteins. For bacteriophage λ, protein affinity chromatography was carried out using the λ N protein,[28] a transcriptional antitermination factor,[29–31] as the ligand. Since *E. coli* strains with mutations in the genes *nusA* (N utilization substance A) and *nusB*[32,33] impaired λ growth due to loss of antitermination, and since antitermination activity of the N protein *in vitro* required the addition of crude extracts from the λ host *E. coli*,[31,34] it seemed likely that host proteins were necessary and that at least one of them might interact directly with the λ N protein.

N protein affinity columns retained only one *E. coli* protein that did not also bind to the agarose bead support. This protein was identified as NusA since the *nusA1* temperature-sensitive mutation[35] caused the identified protein to become temperature-sensitive and altered its charge.[28] When NusA was purified and subsequently immobilized and used for affinity chromatography with extracts from uninfected *E. coli*, it bound only RNA polymerase,[36] indicating that λ N protein was using NusA as an adapter to interact with RNA polymerase. When the other host cofactors, NusB,[32,33] ribosomal protein S10 (RPS10),[37] and NusG[38,39] were used as ligands for affinity chromatography, they each bound another component of this termination/antitermination regulatory system.[40,41] Consistent with these observations, a stable ribonucleoprotein complex that contains N, its host cofactors, and RNA polymerase forms on *nut* site RNA.[42–45] Similar protein affinity chromatography experiments were carried out to identify PPIs among proteins involved in bacteriophage T4 DNA replication[46,47] and recombination.[47,48] For example, when the T4 gene 32 single-stranded DNA-binding protein was immobilized on beads, at least 10 T4 proteins and three *E. coli* host proteins were specifically bound. Nine of the T4 proteins could be identified as being involved in T4 replication and recombination. In 1983, the *E. coli* genes that encoded the host proteins could not be identified—as they could easily be today.

The exquisite specificities of the PPIs observed in these protein affinity chromatography experiments were presumably selected for during evolution, and biologically irrelevant interactions were counter-selected. An important advantage of protein affinity chromatography is that the concentration of the immobilized ligand can be increased so as to detect even quite weak PPIs with affinities less than 10^5 M^{-1}. Weak interactions can also be detected by other methods—e.g., using high concentrations of fusion proteins for pull-downs from cell extracts, overexpressing tagged proteins in cells,[49] or crosslinking proteins in cells prior to their affinity purification (AP)—that are described below.

Limitations of protein affinity chromatography

In the early days, protein affinity chromatography required overcoming two major technical obstacles. The first problem was the requirement for a highly purified protein to be used as ligand. Relatively few proteins had been purified in the 1980s, as purification schemes using conventional column chromatography were often time-consuming and difficult, and purification schemes for specific proteins could take years to develop. Moreover, despite the introduction of recombinant DNA technology in the mid-1970s, relatively few genes had been cloned, vectors for protein overexpression were not yet commonly used, and the use of AP tags to facilitate protein purification had not yet been introduced. Thus, not surprisingly, two other early examples of protein affinity chromatography with crude cell extracts used the highly abundant cytoskeletal proteins actin and tubulin as ligands.[50,51]

The second major obstacle was identifying the interacting proteins. That could be easy if the protein components of the biological system being investigated had already been purified and their mobilities on SDS polyacrylamide gels were known, as was the case for the T4 replication proteins[47] and, eventually, for the proteins involved in transcriptional antitermination by the λ N protein.[38,39] On the other hand, identifying NusA as a host cofactor interacting with λ N protein[28] was possible only because of the fortunate circumstance that NusA was one of the host factors for transcriptional antitermination by λ N protein that had been genetically identified at that time.[35] In the absence of comprehensive genetic investigation, which was rarely the case even in *E. coli* and *Drosophila*, and almost never the case in other organisms, identifying the function of an interacting protein would have had to rely on guessing its function and carrying out an appropriate biochemical assay. This rather intuitive approach led to the purification and identification of the subunits of TFIIF, a general initiation factor that interacts with RNA polymerase II,[52,53] and of a large number of *Drosophila* actin filament and microtubule-binding proteins.[50,51] Identifying the interacting proteins after protein affinity chromatography was likely to require both hard work and either luck or ingenuity in the 1980s.

ADDITIONAL BIOLOGICAL INSIGHTS FROM USING ANTIBODIES AND FUSION PROTEINS TO IDENTIFY PPIs

An alternative to protein affinity chromatography was to produce antibodies against a purified protein and then use those

antibodies to immunoprecipitate the protein from cell extracts together with its interaction partners. In an important early experiment of this type, a host protein of unknown function with an apparent molecular weight on an SDS polyacrylamide gel of 53 kDa,[54,55] a tumor suppressor still known simply as p53, co-immunoprecipitated with SV40 large T antigen, and the tumor suppressor protein RB co-immunoprecipitated with other viral oncoproteins.[56] An important example a decade later, also in cancer biology, was the demonstration that a monoclonal antibody against the viral oncoprotein $pp60^{v\text{-}src}$ and its activated cellular counterpart $pp60^{c\text{-}src}$ co-immunoprecipitated with proteins containing phosphotyrosine from cell extracts and that this depended on the SH2 domain of the *src* proteins.[57] This insight soon led to observations that direct interactions between SH2 domains and peptides containing phosphotyrosine play key roles in cellular signaling networks.[58] Still another important success was the use of a monoclonal antibody against the *Drosophila* TATA box-binding protein (TBP) to co-immunoprecipitate the other subunits of the RNA polymerase II general initiation factor TFIID,[59] known as TBP-associating factors (TAFs), leading to their identification, molecular cloning, and functional characterization as transcriptional coactivators in various species.

As more proteins were cloned during the 1980s, an important advance was the design of fusion proteins in which various partners were joined to the proteins of interest. These fusion partners, or "tags," were peptides or proteins that could easily be purified in a single step by affinity chromatography. The first two such fusion partners, described in 1988, were glutathione-S-transferase (GST) and maltose-binding protein (MBP).[60,61] GST- and MBP-fusion proteins could be purified by binding to glutathione or amylose columns, respectively, followed by elution in native form with glutathione or maltose. Many other protein tags were introduced over the years, including *S. aureus* protein A that binds to immobilized immunoglobulin G (IgG), polyhistidine sequences that bind to nickel or cobalt columns, and small peptide tags for which there were already good antibodies for immunopurification (e.g., influenza virus hemagglutinin [HA] protein). In thousands of such experiments, fusion proteins bound to agarose beads were mixed with cell extracts to "pull down" interacting proteins. In important cancer biology experiments in 1990, recombinant SH2 domains fused to the *E. coli* TrpE protein were bound to beads with immobilized TrpE antibodies and used to pull down epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) receptors that had been phosphorylated on tyrosine.[62,63]

Limitations of using co-immunoprecipitation or pull-downs to identify PPIs

Despite yielding foundational results informing basic and disease biology, co-immunoprecipitation was even less scalable to multi-protein studies than protein affinity chromatography, since it required not only purification to homogeneity of the protein in question but also preparation of a highly specific antibody, which can be very difficult. There was also the possibility that such an antibody might interfere with the assembly of the protein of interest into a protein complex. Also, regardless of the tag used for pull-downs, this method could fail if the fusion partner prevented interaction of the protein of interest with one or more of its partners or prevented its assembly into a protein complex. As well, for many years prior to the advent of sequenced genomes, cloning a gene or cDNA encoding a protein of interest was slow and laborious. It required purification of the protein and then using either an antibody to screen an expression library or oligonucleotides based on peptide sequences to screen a DNA library. Moreover, an overexpressed fusion protein might be soluble in *E. coli*, but often it was not,[64,65] even if it was a prokaryotic protein, and especially if it was a eukaryotic protein that lacked either the post-translational modifications (PTMs), or the chaperones, or both, needed for it to fold properly in its natural host cell. To overcome some of these problems, systems were eventually developed to overexpress proteins in insect, yeast, and mammalian cells, as well as other systems. Regardless of whether protein affinity chromatography, co-immunoprecipitation, or a tag was used, identifying the interacting protein(s) in order to make possible the characterization of its function and biological relevance remained difficult until the advent of genome sequences and protein identification by mass spectrometry (MS).

INSIGHTS FROM USING INTRACELLULAR TAGGING, AP, AND MS TO IDENTIFY PPIs: AP-MS

Introduction of MS to identify interacting proteins

Prior to the determination of genome sequences in the late 1990s, an interacting protein could occasionally be identified if its N-terminal amino acid sequence was determined by Edman degradation[66] and if, fortuitously, its cDNA had already been cloned and sequenced. This picture changed in 1988 when the development of two soft ion desorption methods, electrospray ionization (ESI)[67] and matrix-assisted laser desorption and ionization (MALDI),[68,69] made possible for the first time the determination by MS of the mass-to-charge ratios (m/z values) of large biomolecules, including peptides and proteins. This led in 1993 to the realization that the masses of peptides obtained after digestion of a protein with an amino acid-specific protease like trypsin could be determined by MS and then used to search a DNA or amino acid sequence database in order to identify the protein.[70,71] The sequencing of many genomes less than a decade later then ushered in the current era of "proteomics" in which it became possible to associate large numbers of proteins with specific gene sequences—and even with specific functions to the extent that mutations in those genes had been shown to have specific phenotypic consequences. An important improvement that provided much more identifying information about individual peptides in complex mixtures and thereby greatly facilitated the identification of proteins derived from large genomes was the development of liquid chromatography-tandem mass spectrometry (LC-MS/MS), in which peptides were separated by LC and then fragmented by collision-induced dissociation (CID) in the mass spectrometer,[72] as well as algorithms that facilitated automated searches of sequence databases. Since then, many improvements in design have improved the sensitivity and mass accuracy of mass spectrometers.

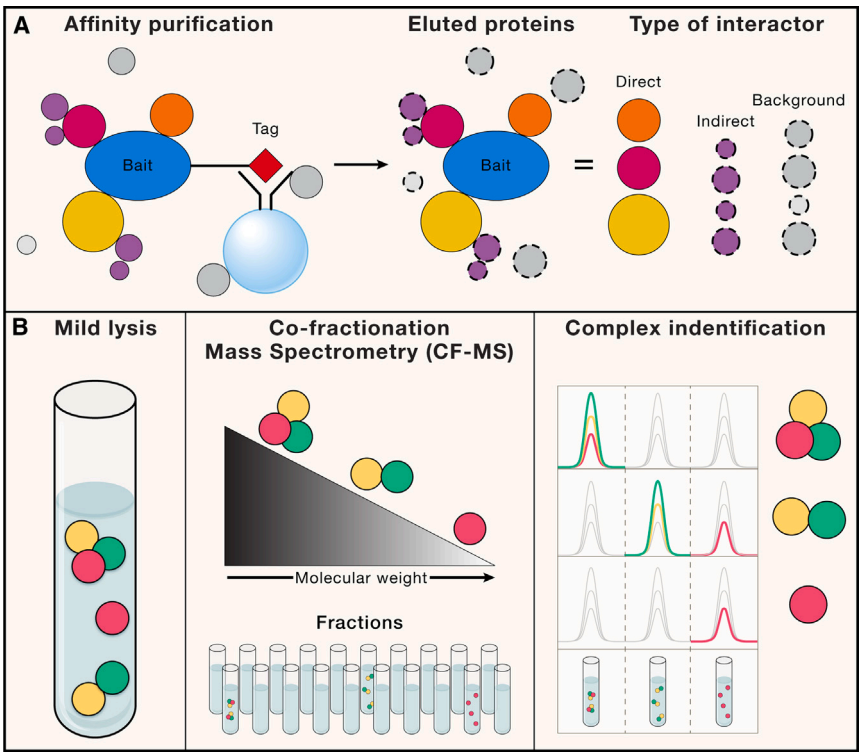

Figure 2. Methods for detecting PPIs
(A) Affinity purification and mass spectrometry (AP-MS).
(B) CF-MS (co-fractionation-mass spectrometry). Adapted from Richards et al.[73]

Intracellular protein tagging followed by AP

Beyond advances in instrumentation, there were important advantages to be had by expressing a tagged protein in its natural host cells prior to the preparation of cell extracts, then using an immobilized reagent that bound the tag to purify the tagged protein and its interacting proteins for identification by MS. First, using the natural host usually avoided the common problem of protein insolubility that often occurred when a protein was overexpressed in a heterologous system. Second, the tagged protein would acquire its natural PTMs, and these might contribute to protein solubility and/or the association of certain interacting proteins. Third, the protein would be synthesized in the presence of its natural chaperones and co-chaperones, which might be necessary for it to fold properly. Finally, if the DNA encoding the tag was integrated into the gene that normally encodes the protein so as to label the protein's N or C terminus, the tagged protein would be expressed at its normal level—thereby potentially avoiding artifactual interactions caused by protein overexpression—although protein overexpression might be needed in some cases to prevent dissociation and detect weak interactions. Combined with protein identification by MS, this procedure came to be known as AP and MS (AP-MS) (Figure 2A).

Biological insights from AP-MS

Because intrachromosomal protein tagging in the budding yeast *S. cerevisiae*,[74] and eventually, *E. coli*,[75,76] was relatively easy to do, it became possible to carry out AP-MS for essentially all the proteins of these organisms. Using the "tandem affinity purification (TAP) tag" (which consists of two tags separated by a tobacco etch virus [TEV] cleavage site) and two sequential AP steps,[77] one could readily purify to near homogeneity both the tagged proteins and their interactors. This approach made possible the systematic, genome-wide identification of PPIs and protein complexes from both yeast and *E. coli*.[78–82]

The presence of various proteins in the same protein complex implies common, or at least overlapping, functions via "guilt-by-association." Thus, identification of the yeast protein complexes then led to a large number of insights into yeast chromatin biology and various other aspects of its molecular biology (e.g., Bouveret et al.,[83] Miller et al.,[84] Pijnappel et al.,[85] Krogan et al.,[86] Krogan et al.,[87] Dziembowski et al.,[88] Krogan et al.,[89] Kim et al.,[90] and Keogh et al.[91,92]). For example, insights were obtained into histone methylation by COMPASS (complex proteins associated with Set1), the first histone methyltransferase complex to be purified,[84,86] and histone deacetylation by the SET3 complex.[85] Another discovery was that there are two histone deacetylase complexes containing Rpd3, one that functions at promoters and another that functions in transcribed regions.[92] Yet another advance came with identifying that there is a specialized chromatin remodeling complex that replaces histone H2A with its variant Htz1.[87] Other studies offered insights into the control of DNA repair[89,91] and termination of transcription by RNA polymerase II.[90]

Today the power of AP coupled with modern MS and its accompanying bioinformatic procedures are such that a single AP step usually suffices for the unambiguous identification of metazoan peptides and, therefore, of a tagged protein and its various interactors. Purifying a large number of different proteins fused to the same tag creates a relatively uniform background of contaminating proteins that had either bound to the tag or the matrix used for the purifications. Also, reference to a database of the common contaminants in APs[93] (in conjunction with statistical algorithms like SAINTexpress,[94] CompPASS,[95] and MiST[96]) usually makes it possible to discriminate between true interactors and various contaminants. Using these or similar methods, metazoan PPI networks based on AP-MS were generated for ~5,000 *Drosophila* proteins[97] and for up to ~5,000 protein-encoding cDNAs from the ORFeome collection[98] for human cell lines.[99,100] These and many other human PPIs and protein complexes derived from various smaller-scale studies have been compiled in various databases like CORUM, InWeb_IM, and BioGRID.[101–103]

As a result, a very large number of human protein complexes or "communities" have been resolved, which now encompass

the majority of human proteins. Moreover, the AP-MS procedure can be used to compare the effects on PPIs of different cell types, various perturbations, and mutations that cause disease. For example, large-scale AP-MS experiments have increased our understanding of signaling pathways, some of which can be directly connected to disease states. The mapping of PPIs for 300 human kinases identified ~7,000 interactions across 2,379 unique proteins.[104] This dataset provided insight into poorly understood kinases, including linking the kinases PIM3 and POMK to cytoskeleton regulation and glycoprotein processing, respectively, and also made previously unrecognized connections between several kinases and a number of disease areas. AP-MS studies have also shed light on the importance of PTMs in mediating PPIs, including efforts to explore interactions involving proteins in the Hippo signaling pathway,[105] which regulates various aspects of cell proliferation and immune system function and has been implicated in cancer, as well as cardiovascular system function and disease.[106] Identification of PPIs after treating extracts with phosphatase inhibitors revealed phosphorylation-dependent interactions mediated by the STRIPAK kinase complex and the MST1/2 kinases of the Hippo pathway.[105] Also, AP-MS experiments carried out after the addition of phosphatase inhibitors uncovered additional differential interactions involving MOB1A, a key scaffold protein of the system.[105] For additional discussion of such differential physical interactions, see below.

Limitations of PPI identification by AP-MS

As mentioned above, an important limitation of AP-MS for the identification of human PPIs is that some may be artifactually driven by the overexpression of the tagged protein. In fact, networks generated by intrachromosomal tagging for *S. cerevisiae* and *E. coli*[78–82] do not overlap very well with those generated for both organisms after protein overexpression.[49,107] Protein overexpression can be avoided for human proteins by using CRISPR technology for intrachromosomal tagging, and this has been done for a limited set of human proteins.[108] On the other hand, some interactions that have *in vivo* relevance may be so weak that they dissociate during the washing steps involved in AP, and some of these may be retained by protein overexpression. Another issue that is common to all methods that use whole-cell extracts, including protein affinity chromatography, pull-downs, immunoprecipitations, and AP-MS, is the possibility that some proteins may interact in whole-cell extracts, even though they can never interact *in vivo* because they are always localized in different cellular compartments or are never expressed at the same time. In the absence, however, of pre-existing genetic data—e.g., from genome-wide association studies (GWASs)—only laborious, low-throughput, biochemical and genetic methods are currently available to decipher whether a PPI discovered as a consequence of protein overexpression or the use of cell extracts is biologically meaningful.

PPIs vary substantially among the many specialized cell types that are present in metazoans. Moreover, they can also vary during development, in response to environmental stimuli, and at various points in the cell cycle, and it would be extraordinarily arduous to carry out genome-wide tagging followed by AP-MS in all these circumstances. A solution to this issue is discussed in the following section of this review. Another issue is the challenge of membrane proteins, which are not soluble in the buffers commonly used to prepare cell extracts. To address this, mild detergents have been used that can solubilize membrane proteins without disrupting protein complexes.[82] Still another issue that is common to AP-MS and most other biochemical methods used to identify PPIs is the question of whether an observed PPI is direct or is instead indirectly mediated by another protein in the same protein complex (Figure 2A); this issue can also be resolved in ways discussed below.

PROTEIN CO-FRACTIONATION FOLLOWED BY MS TO CHARACTERIZE PROTEIN COMPLEXES

In early protein purifications, polypeptides were thought likely to be components of the same protein complex if they co-chromatographed in a series of purification steps. This principle was utilized in the development of co-fractionation-MS (CF-MS), in which the proteins present in the fractions resulting from column chromatography are digested with trypsin, then identified and quantified by tandem MS (Figure 2B).[109,110] The fractionations can be carried out using high-resolution ion exchange chromatography, size exclusion chromatography, isoelectric focusing, or hydrophobic interaction chromatography—and by combinations of these methods. Algorithms then detect co-elutions for all the proteins that can be identified by MS, as many as 8,389 in the case of mouse brain.[111] Co-eluting polypeptides tentatively assigned to the same protein complex suggested the subunit compositions of as many as 1,030 mouse brain protein complexes, nearly 40% of which appeared to be new.[111] For example, a large and novel protein complex contained multiple proteins linked to amyotrophic lateral sclerosis (ALS), appears to have a role in splicing, and is altered in a mouse model of ALS.[111] Other novel protein complexes contained proteins linked to various other neuropsychiatric conditions.

Similarly, a study looking at complexes in jumbophage-infected bacteria discovered large numbers of new phage-phage and phage-bacteria PPIs related to jumbophage predation.[112] CF-MS is particularly well suited for the study of cells infected by pathogens because it avoids the need for tagging and purifying all the proteins of both the host and the pathogen for MS analysis. Moreover, it can examine not only pathogen-pathogen and pathogen-host interactions but also the indirect effects of the pathogen on the PPI network and protein complexes of the host.

Advantages and limitations of CF-MS

Because the co-purification of two proteins may be fortuitous, it is important to validate the PPIs and protein complexes identified by CF-MS by comparison to known PPIs and protein complexes. For example, 18,774 of the 25,235 PPIs as well as 498 of the 772 protein complexes identified by CF-MS in human breast cancer and non-transformed cell lines[113] were supported by PPIs previously described in the literature or the protein complexes present in various databases. This left 274 hypothetical new protein complexes that remain to be validated by other approaches. Likewise, because the identification of a PPI by AP-MS may be artifactually driven by protein overexpression, it will also be important to use CF-MS to validate PPIs identified by AP-MS. Therefore, AP-MS and CF-MS should be highly complementary, and the systematic pursuit of both types of studies

will hopefully lead eventually to highly reliable descriptions of the compositions of human protein complexes and how they vary among different cell states.

An important limitation of CF-MS is the inability of MS to detect all the proteins present in a cell or tissue, a limitation that may be resolved as MS technology improves. In addition, in all CF-MS studies, a single protein is sometimes found in more than one peak in various chromatograms. This could reflect either its presence in more than one protein complex or else the partial dissociation of a protein complex into multiple sub-species during column fractionation, an issue that could also be resolved by AP-MS. Unlike AP-MS, however, the CF-MS method has the advantage that neither tags nor antibodies are used that could disrupt protein complexes. It also has the advantage that different cell types or treatments can be compared quantitatively by employing mass tags to differentially label proteins in different cell populations prior to mixing the extracts for column chromatography and MS, as in the SILAC[114] method, or by chromatographing the extracts separately and using a data-independent acquisition (DIA) MS method like SWATH-MS (sequential window acquisition of all theroetical mass spectra).[115] The throughput of CF-MS can be increased substantially by using 6-plex,[113] or even 18-plex, tandem mass tags (TMTs),[116] thereby multiplexing the MS analyses and simultaneously enabling quantitative comparison of multiple samples and reducing instrument time.

VALIDATION OF PPIs AND DISTINGUISHING DIRECT FROM INDIRECT INTERACTIONS

Validation of PPIs

If particular PPIs identified by high-throughput methods like AP-MS or CF-MS are to be exploited for further research or the development of therapeutics, they need to be validated. The detected PPIs are much more likely to be valid if genetic studies have indicated that the interacting partners are components of the same biological system and if the interacting partners are present in the same intracellular compartment. They are also more likely to be valid if their expression is correlated across cell types and cell perturbations, if their thermal stabilities in cell extracts are highly correlated,[117] and especially if they are validated by an independent method.

A number of methods have been developed for the direct validation of particular PPIs. The most commonly used involves IP-Westerns, where an interacting prey protein is detected by immunoblotting after pulling down a bait protein using an immobilized antibody against either the bait protein or a tag. This method does not eliminate the possibility that a PPI is either indirect, involving an additional protein or nucleic acid to support the interaction, or artifactual, where the interaction only occurs *in vitro* after cell lysis.

Other methods have been developed to determine whether a proposed interaction occurs *in vivo*. One such method involves resonance energy transfer (RET) *in vivo* from a fluorescent (FRET) or bioluminescent (BRET) donor to which the bait protein is fused to a FRET acceptor to which the prey protein is fused. Light emission from the acceptor after stimulation of the donor then depends on the physical distance between the acceptor and the donor and, therefore, on the PPI.[118–120] Another approach involves proximity ligation assays (PLAs) that are conducted in fixed cells. Putatively interacting proteins are detected by antibodies, each of which is fused to an oligonucleotide. When the proteins, and therefore the antibodies, are in proximity, a single-stranded DNA molecule can bridge the oligonucleotides and create a substrate for detection of the interaction by rolling circle amplification.[121,122] Other important innovations have been the development of proximity labeling methods like those using the engineered biotin ligase-based BioID and TurboID, or engineered ascorbate peroxidase (APEX) to detect potential PPIs *in vivo*.[123–125] These methods rely on the expression of a fusion protein in which a bait protein is fused to an enzyme that can be stimulated to cause the biotinylation of nearby proteins. The biotinylated proteins can then be purified using streptavidin beads and identified by MS, and these methods have been extensively utilized to identify the proteins present in various cellular compartments and sub-compartments.[126] However, the proteins that become labeled with biotin can be sufficiently far away that they are not necessarily components of the same protein complex. More recently, a sophisticated method was described that exploits protein *trans*-splicing with a split intein to install a photo-activatable catalyst into a bait protein; this allows shorter-range proximity labeling with biotin in isolated nuclei, and it is more likely to detect partners in the same protein complex.[127] However, none of these methods can distinguish whether a PPI is direct or mediated by another component of the same protein complex.

Distinguishing direct from indirect interactions

In view of the possibility of developing drugs that function by enhancing or interfering with PPIs that have been affected by disease-related mutations (see below), it will often be important to know whether PPIs identified by AP-MS or CF-MS are direct or mediated by an intermediary protein or even a nucleic acid. Most obviously, recombinant versions of the proteins could be prepared and tested for interaction *in vitro*. In addition to biochemistry, several genetic systems have been developed to identify binary PPIs, the first being the yeast two-hybrid (Y2H) system.[128] In the Y2H system, two proteins are expressed in *S. cerevisiae*, one protein fused to the DNA-binding domain of the yeast activator protein Gal4, and the other fused to the activation domain of Gal4. Interaction of the two proteins of interest leads to activation of one or more reporter genes containing upstream Gal4-binding sites. Various improvements in the Y2H framework then made it possible to screen for millions of potential PPIs[129,130] (e.g., for human proteins, where the potential number of protein pairs to be screened, if one ignores splice variants, is approximately 2×10^8). Ultimately, cDNAs encoding 17,408 different human proteins were screened against each other,[131] resulting in the identification of 56,406 binary PPIs involving 9,094 proteins that were validated by orthogonal binary interaction assays. This corresponds to ~2%–11% of the binary human PPIs that are thought to exist.[131] An earlier Y2H effort covering about half of that search space recovered ~27% of the PPIs that were found in multimeric structures in the Protein Data Bank.[132] However, the Y2H method is less likely to be successful at detecting the many PPIs in large protein complexes, and the Protein Data Bank was strongly biased toward the structures of smaller protein complexes (e.g., dimers) rather than large

protein complexes. This trend is changing as cryogenic electron microscopy (cryo-EM) is used to produce more and more high-resolution structures of protein complexes.

Beginning with the split ubiquitin system,[133] many other binary systems have been developed[134] based on split proteins whose two segments are brought together for readout by the interaction of two proteins to which the segments of the split protein had been fused. One limitation of the Y2H system is that the PPIs must occur in the nucleus, whereas the PPIs of the split protein systems could, in principle, be localized to other cell compartments. In fact, one version of the split ubiquitin system was designed specifically to detect PPIs localized to the cell membrane.[135] While useful, a significant fraction of the PPIs may be missed with the Y2H and split protein systems because the fusion partners may interfere with either the PPI or the incorporation of one of the proteins into a complex. Also, some of the PPIs may be missed if the interacting proteins need chaperones or cell-type specific PTMs to bind. Finally, the binary methods may often be unable to detect PPIs in large protein complexes stabilized by multiple weak PPIs.

With an eye on exploiting direct PPIs therapeutically, the ability to visualize the interacting surfaces would also be key. The polypeptide chains in close proximity in a large protein complex isolated by AP can now be identified quite reliably using crosslinking MS (XL-MS).[136,137] In this procedure, the isolated protein complex is treated with a bifunctional chemical crosslinking reagent, most often one that targets lysines or carboxyl groups. Following digestion with trypsin and often AP of the crosslinked peptides, the sequences of crosslinked peptides are determined by LC-MS/MS. This procedure is commonly used as an adjunct validation method for structure determinations by cryo-EM[138] (see below). Importantly, if cells are treated with the crosslinking reagent prior to protein complex isolation by AP, the crosslinks are indicative of PPIs that occur *in vivo*.[139] Moreover, the use of this method can also make it possible to identify weak PPIs that cannot survive the washes needed for AP. Finally, XL-MS should help scientists to identify the surfaces mediating contact between the two proteins, which will be important for the development of drugs designed to disrupt that interface. Development of methods for the tag-free identification of crosslinked peptides on a proteome-wide scale is an active area of investigation.

USE OF AP-MS WITH WILD-TYPE AND MUTANT HUMAN AND VIRAL PROTEINS FOR INSIGHTS INTO DISEASE

An emerging frontier is assessment of temporal and spatial changes[140] in PPIs. One would like to know how protein interactions are altered in different parts of the cell[141] and how the interactions change as a function of time when conditions change.[142,143] For example, recent studies focused on PPI changes during the response to DNA damage[142,143] and during stem cell differentiation.[144] It is also possible to uncover the effects of specific mutations associated with disease states. For example, PPI networks have been generated using protein variants associated with neurodegenerative diseases, including Alzheimer's disease (AD), Huntington's disease, and Parkinson's disease.[145] Many proteins were shown to display differential binding when comparing a mutant and its corresponding wild-type protein. For example, a mutant form of amyloid precursor protein (APP), K670N/M671L, which is associated with early-onset AD, is altered in its interaction with LRPPRC, an essential regulator of the mitochondrial respiratory chain. This change was reported to induce mitochondrial dysfunction, a phenotype that correlates with AD,[145] demonstrating that this approach can shed light on the biology underlying neurologic disorders.

Differential AP-MS experiments have also been carried out for many cancer-causing mutations, including RAS gain-of-function missense mutations[146] and PP2A phosphatase mutants.[147] Recent work has systematically analyzed the effects on PPIs of the most frequent mutations associated with breast and head/neck cancer in multiple cell types.[148–150] This uncovered a myriad of mutant- and cell-type-specific PPIs involving key oncogenic drivers, including BRCA1 and PIK3CA. For example, two of the most prevalent head/neck cancer mutations fall in the helical domain of PIK3CA (E542K and E545K) and result in this enzyme's increased affinity for the upstream-acting tyrosine kinase, HER3. Interestingly, a tighter physical connection between PIK3CA and HER3 seems to correlate with a stronger response to HER3 inhibition *in vivo*,[148] suggesting that data from PPI maps could have clinical implications.

A similar systematic AP-MS approach was used on a set of ~100 autism spectrum disorder (ASD) risk genes,[151] revealing 1,800 PPIs, the vast majority of which had not been previously reported. A total of 54 patient-derived missense variants were introduced into specific proteins for subsequent AP-MS analysis, which created an ASD mutant differential interaction map after comparison to wild-type. One mutation in the transcription factor FOXP1, which resulted in loss of interaction with its binding partner FOXP4, led to a number of neuronal phenotypes in brain organoids, including an increase of subplate neurons, consistent with what is known about ASD biology.[151] These results suggest differential PPI maps can be used to prioritize disease mutations that may represent actionable therapeutic targets (see below).

In both sets of studies,[148–151] convergence was also observed in the sense that the initial set of variant genes could be reduced to a smaller number of complexes and pathways since a high degree of connectivity was observed among the corresponding proteins, thus demonstrating how PPI maps can aid in interpreting disease gene discovery efforts. In fact, it would be beneficial to have gene discovery analysis carried out not at the individual gene level but using protein complexes and functional pathways where this biological convergence could be detected (Figure 3A).

Finally, differential AP-MS experiments can also be used to assess the effects of mutations associated with viral variants. For example, systematic studies of the individual mutations seen among the major SARS-CoV-2 variants revealed insights into how the virus was evolving to overcome cellular defense mechanisms, as well as treatments, and thus to effectively hijack the host during infection.[152,153]

BIOCHEMICAL RESOLUTION AND FUNCTIONAL UNDERSTANDING OF PROTEIN MACHINES AND COMPLEXES

We can now envision that, with time and appropriate resources, all stably assembled protein complexes can and will be isolated by AP, and their protein compositions identified by MS. Although

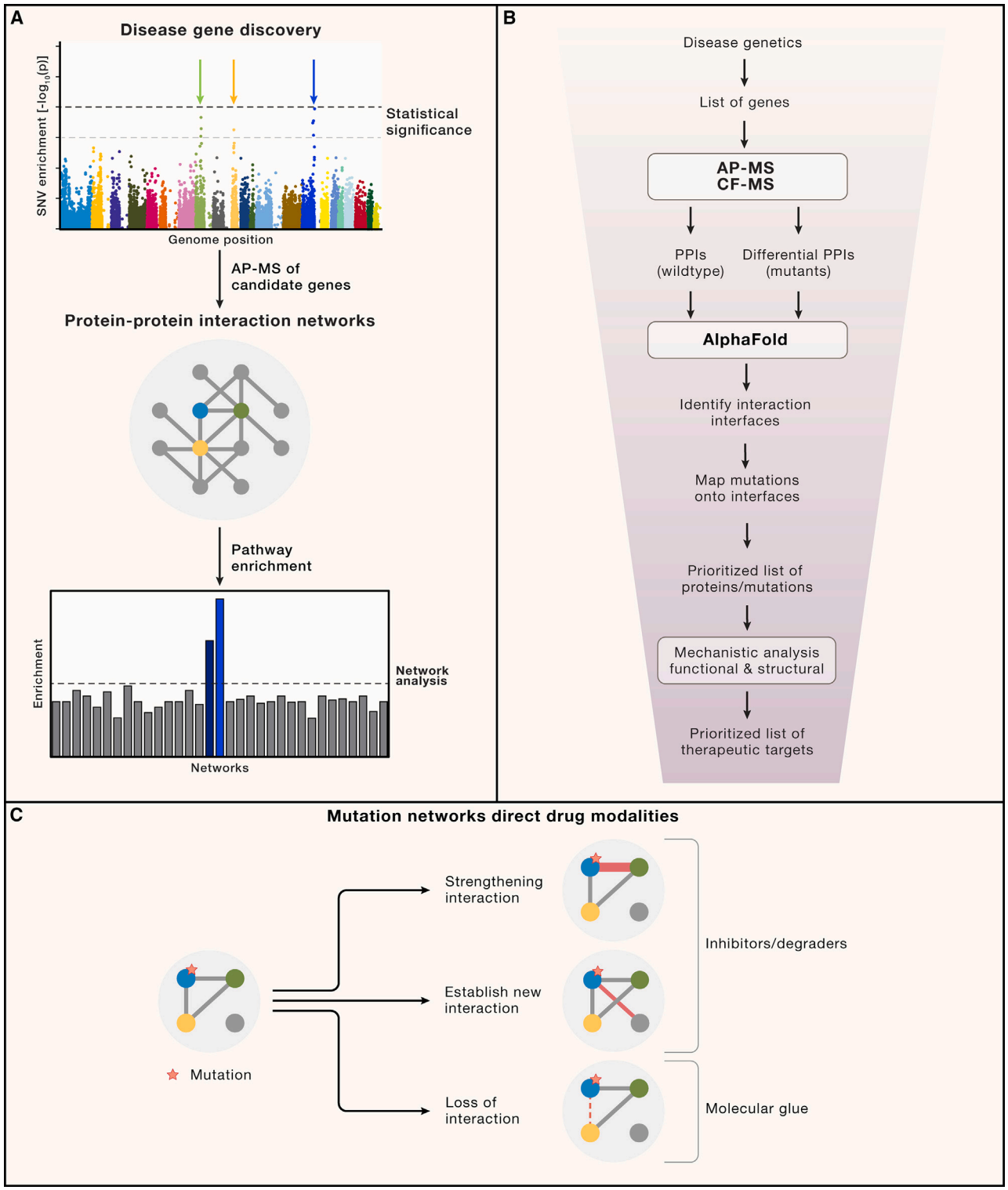

Figure 3. Using PPI maps to interpret disease genetics and provide therapeutically valuable insights

(A) Although few genes are typically able to pass genome-wide levels of significance in GWAS, PPI data can help interpret the results from disease gene discovery efforts. Physical interaction maps of multiple genetic variants, each of which is observed rarely in a disease population, are often found to impact a common region of the network, highlighting significantly affected pathways.

(B) A proposed proteomic, genetic, and structural pipeline to analyze mutations linked to disease. Adapted from Wang et al.[151]

(C) Differential mutant PPI networks can inform drug modality and therapeutic approaches.

the subunit composition of a protein complex and even its function can vary as a function of cell type, CF-MS is capable of revealing the subunit compositions characteristic of particular cell types—and thus the particular cell type that should be used for the purification of each version of the protein complex. However, protein complexes often exist in multiple forms. Even in the same cell, there can be multiple versions of a protein complex with different but overlapping subunit compositions. One example of this is the presence of at least two different histone deacetylase complexes containing the catalytic subunit Rpd3 in the budding yeast *S. cerevisiae*.[92,154] One of these, the larger complex Rpd3L, represses transcription initiation at specific promoters, and the other, the smaller complex Rpd3S, functions as part of a pathway that suppresses spurious transcription initiation in transcribed regions. In one study, the two complexes were resolved by conventional biochemical fractionation,[154] while in another, the complexes were resolved by placing AP tags on subunits unique to either complex.[92] These studies revealed that complementary biochemical approaches can be used to decipher the composition of complexes that contain overlapping proteins. A much more complicated version of this type of application is illustrated by the human SWI/SNF chromatin remodeling complexes (BAF complexes), which are often mutated in cancer[155,156] and neurodevelopmental disorders.[157] There are many different versions of the BAF complex,[156] yet they can be resolved from one another using not only biochemical but structural approaches.[158]

Obtaining a list of protein complexes from a defined set of genes is crucial, but a major next step is, of course, to understand the function of the complexes. How to decide which complexes to focus on, when so many remain mysterious? Much help can come from knowledge of the cell types and compartments in which a particular protein complex is present and from the effects of mutations in components of those complexes on molecular, cellular, and organismal phenotypes. However, when such data are missing, an argument can be made for prioritizing those complexes that are suspected to function as protein machines. Protein machines are often formed as a complex of ten or more gene products. Central to their function is the harnessing of energy to drive unidirectional conformational changes, which can allow the complexes to serve as motors, clocks, and assembly factors.[1] While cells contain many well-studied molecular machines, there are hundreds whose structure and function are not known. Because the free energy change required to create a directional allosteric change in the conformation of the proteins often comes from the hydrolysis of a high energy bond (e.g., ATP or GTP), conserved nucleotide binding sites can help identify protein complexes that can be suspected to exhibit machine-like behavior. Such a protein complex would be expected to transition between sets of different conformations, depending on whether ATP, ADP +P_i, ADP, or no nucleotide is bound.

The standard way of dissecting how each protein machine functions often requires reconstitution of the system in which it works with purified components, a process greatly facilitated by the use of tags for protein purification. This then allows a detailed analysis of the biochemistry of the machine's reactions, using a variety of kinetic and chemical analyses following perturbations of its different components. But the expected transitions between different conformations make protein machines prime targets for the increasingly powerful cryo-EM technologies. This approach has previously been used to analyze the transitions between multiple states of a protein complex through painstakingly difficult experimentation.[159]Today, once PPI experiments have identified a complex of special interest, it can often be captured by AP directly onto cryo-EM grids.[160] Other advances in cryo-EM technologies may now make it possible to determine the "conformational landscape" of a protein complex from the set of different conformations that are present in a mixture.[161] Manipulating ATP and the ligands that mimic various bound nucleotide states increases the information that can be gained from these landscapes for a novel protein machine. Classical protein affinity chromatography techniques[27] can also be used to isolate and screen for the effects of weakly bound "accessory proteins" in such structural analyses. Thus, recent breakthroughs in cryo-EM techniques promise to greatly speed our understanding of how machine-like protein complexes function. Finally, one can hope that improvements in the techniques that allow different conformations of a protein complex to be determined in slices of whole cells using cryogenic electron tomography (cryo-ET)[162,163] can be systematically applied to all complexes to determine not only their locations in the cell but also their different conformations and ultimately functions.

HOW AI AND CRYO-EM, IN CONJUNCTION WITH AP-MS AND CF-MS, CAN IMPACT THE DEVELOPMENT OF PPI-DIRECTED THERAPIES

In order to extract biological and therapeutic value from differential PPI datasets, it will be valuable to know which mutations that perturb PPIs do so by directly affecting a specific interface. However, one major caveat with data derived from AP-MS studies has always been that it is difficult to determine which PPIs reflect direct interactions compared with those that are indirect and mediated via other proteins (Figure 2). Cross-link analyses by XL-MS and structure determination using cryo-EM are both capable of resolving this issue, but these methods are either laborious (cryo-EM) or lack comprehensiveness (XL-MS), especially considering the need to assess the effects of very large numbers of disease-causing mutations on PPIs. Some of these issues would be resolved if innovations in experimental and computational methods make it possible in the future to identify all, or almost all, the crosslinked peptides, even the ones with relatively low abundance.

Fortunately, recent advances in protein structural predictions using AI-based methods are beginning to have a profound impact on understanding and prioritizing PPIs. Most notably, AlphaFold[164–166] has become increasingly accurate at predicting protein structures, and these methods are being extended to predict PPIs with increasing success.[167–169] For example, by using AlphaFold to screen for PPIs among proteins linked to DNA metabolism, Lim et al.[170] identified the previously unknown DONSON complex and then demonstrated that it is required for efficient initiation of eukaryotic DNA replication. Furthermore, AlphaFold has been used to help predict which

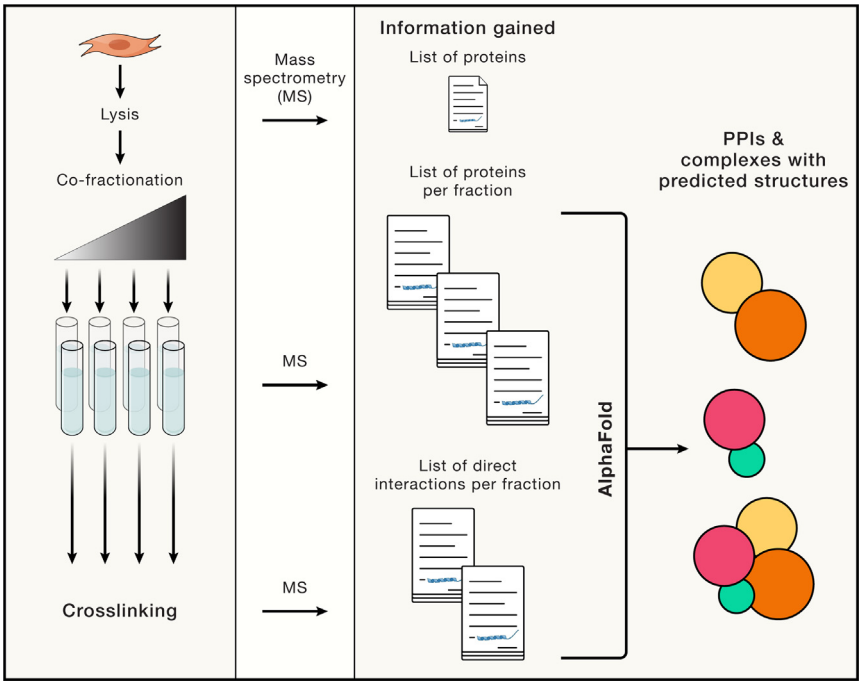

Figure 4. Vision for the generation of global, high-resolution PPI maps using CF, XL-MS, and AI approaches

Column fractionation following cell lysis, combined with XL-MS, would provide a global, albeit lower resolution, view of PPIs and protein complexes. Higher resolution would be achieved when AI approaches, including AlphaFold, are used on the resulting datasets generated at multiple points in the pipeline.

of the AP-MS-derived PPIs involving proteins connected to autism are direct, thereby helping to identify novel proteins connected to the disease state.[151] Furthermore, identifying those disease mutations that reside in the interfaces of the direct PPIs predicted by AlphaFold will allow for prioritization of the interactions and mutations that should be studied in greater detail using targeted genetic studies and structure-based efforts using cryo-EM (Figure 3B).[151]

This integrative approach should not only identify new targets for the development of drugs to treat genetically defined diseases but also give insight into how to "target" the targets. For example, a mutation that results in the introduction of either a new interaction or the strengthening of a previously existing one (e.g., E542K or E545K mutations of PIK3CA that strengthen its interaction with HER3; see above) suggests that inhibiting the function of the interacting partner may represent an effective therapeutic strategy. On the other hand, if a mutation perturbs an interaction, resurrecting the PPI, potentially through molecular glues, may be advantageous (Figure 3C). As recently reviewed by others, PPI data can, in general, be leveraged therapeutically through bifunctional proteolysis targeting chimeras (PROTACs) and molecular glue degraders.[171] In addition, the development of AI-based approaches that accurately predict what compounds would interact specifically with PPI interfaces may prove valuable.

How AI and XL-MS may revolutionize the genome-wide identification of PPIs in the future

In the future, it would be advantageous to quickly and systematically generate high-resolution, comprehensive PPI maps in an organism or under different conditions, including in the presence of specific stresses, drugs, or disease-causing mutations, and perhaps even accomplishing this feat at single-cell resolution. While this is presently not possible, future developments in AI and XL-MS technologies may make it possible to approach these goals. For example, one could envision further development of the methods based on AlphaFold that were used to dock pairs of interacting proteins[166–169] being used, first to identify all human proteins predicted to interact with one bait protein, then being extended to search for all possible pairwise interactions. Of course, the computational cost of evaluating all of the $\sim 2 \times 10^8$ possible human pairwise PPIs would, at least now, be enormous, and the results would be blind, at least initially, to the effects of alternative splicing and the very large number of PTMs that regulate PPIs. Moreover, it would be difficult to also account for possible cooperative interactions involving more than two proteins within a protein complex. Even with these important caveats, this kind of information, even now, is enabling significant reduction of time and resources as these computational approaches are effectively prioritizing direct and functionally relevant PPIs.

A presently more feasible approach to achieve this goal would be to carry out CF-MS followed by XL-MS on each of the column fractions to provide an overview of the PPI landscape in extracts from various types of cells. Such an approach would deal with the issue of non-interacting proteins that fortuitously happen to co-fractionate, and it would also provide information about the topologies of the interacting proteins. Combining this experimental analysis with AI approaches, including AlphaFold, to help refine the XL-MS-derived datasets could, in theory, provide a comprehensive view of the PPI landscape (Figure 4). Analysis by MS is becoming increasingly sensitive (including to single-cell resolution),[172] and so, in theory, this approach could even be carried out at single-cell resolution, but the sensitivity of the relevant instrumentation would need to be substantially increased. Even if this approach provided a lower resolution view, the data could aid in identifying and prioritizing key complexes and insights that could inform more targeted functional and structural studies.

Excitingly, methods are being developed to use XL-MS to identify crosslinked peptides in isolated organelles[173] or even in whole cells after treatment with a membrane-permeable crosslinking reagent.[174,175] These approaches should overcome many of the key limitations related to spurious binding in extracts and failure to capture transient interactions. Such methods can identify protein topologies as well as PPIs, but whether they

can be made sensitive enough to identify nearly all of the PPIs in a whole cell or even in an isolated organelle remains to be seen.

Conclusions

In the past, attempts have been made to glean both biological and therapeutic insights directly from genes and mutations derived from disease-focused genomic sequencing efforts. However, simply put, there is a limit to the insights that can be extracted from sequence information, and further sequencing efforts on specific disease areas often have diminishing returns. Importantly, the functional units of the cell are proteins, and they often talk to one another through physical interactions. These connections, and the molecular machines the proteins ultimately form, are involved in virtually all functions of the cell. Furthermore, mutations associated with disease often perturb the function of proteins, as well as their interactions, and the vast majority of therapies that are developed to treat disease target proteins. Therefore, to understand healthy biology and ultimately treat disease states, a deep and mechanistic understanding of PPIs is crucial. While the study of PPIs has existed for almost half a century, exciting recent advances in the integration of experimental and computational approaches have provided unprecedented ways to identify and characterize homeostatic PPIs and the ones affected by disease-causing mutations through a high-resolution lens that has not previously existed. These views, emerging from data-rich, integrated pipelines, will usher in a new era of identifying and understanding PPIs that will be exploited to more deeply understand the molecular biology of health and disease as well as to develop new therapeutic approaches for a myriad of diseases.

ACKNOWLEDGMENTS

The authors would like to thank Manon Eckhardt for help with scientific editing and figure design, Alicia Richards for help with the literature review, and David Agard, Matthew State, George Scangos, and Kevan Shokat for their invaluable discussions. This work was supported by grants from the National Institutes of Health (U19AI135990, U54AI170792, U54CA274502, 1OT2OD032742, and U19AI135972 to N.J.K.) and the Canadian Institutes of Health Research (FDN-154338 to J.F.G.). This work was supported by the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement #HR-0011-19-2-0020 and #HR-0011-20-2-0029. The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

DECLARATION OF INTERESTS

The Krogan Laboratory has received research support from Vir Biotechnology, F. Hoffmann-La Roche, Maze Therapeutics, and Rezo Therapeutics. N.J.K. is on the Board of Directors of Rezo Therapeutics, and he is a shareholder in Tenaya Therapeutics, Maze Therapeutics, Rezo Therapeutics, GEn1E Lifesciences, and Interline Therapeutics.

REFERENCES

1. Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*, 291–294. https://doi.org/10.1016/s0092-8674(00)80922-8.
2. Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. Science *292*, 1863–1876. https://doi.org/10.1126/science.1059493.
3. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature *560*, 607–612. https://doi.org/10.1038/s41586-018-0440-4.
4. Brautigam, C.A., and Steitz, T.A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. Curr. Opin. Struct. Biol. *8*, 54–63. https://doi.org/10.1016/s0959-440x(98)80010-9.
5. Budenholzer, L., Cheng, C.L., Li, Y., and Hochstrasser, M. (2017). Proteasome Structure and Assembly. J. Mol. Biol. *429*, 3500–3524. https://doi.org/10.1016/j.jmb.2017.05.027.
6. Bard, J.A.M., Goodall, E.A., Greene, E.R., Jonsson, E., Dong, K.C., and Martin, A. (2018). Structure and Function of the 26S Proteasome. Annu. Rev. Biochem. *87*, 697–724. https://doi.org/10.1146/annurev-biochem-062917-011931.
7. Ye, Y., Wu, H., Chen, K., Clapier, C.R., Verma, N., Zhang, W., Deng, H., Cairns, B.R., Gao, N., and Chen, Z. (2019). Structure of the RSC complex bound to the nucleosome. Science *366*, 838–843. https://doi.org/10.1126/science.aay0033.
8. Hang, J., Wan, R., Yan, C., and Shi, Y. (2015). Structural basis of pre-mRNA splicing. Science *349*, 1191–1198. https://doi.org/10.1126/science.aac8159.
9. Zhang, X., Yan, C., Hang, J., Finci, L.I., Lei, J., and Shi, Y. (2017). An Atomic Structure of the Human Spliceosome. Cell *169*, 918–929.e14. https://doi.org/10.1016/j.cell.2017.04.033.
10. Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science *289*, 905–920. https://doi.org/10.1126/science.289.5481.905.
11. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., et al. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. Cell *102*, 615–623. https://doi.org/10.1016/s0092-8674(00)00084-2.
12. Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. Nature *407*, 327–339. https://doi.org/10.1038/35030006.
13. Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 A resolution. Science *292*, 883–896. https://doi.org/10.1126/science.1060089.
14. Hutchison, C.A., 3rd, Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., et al. (2016). Design and synthesis of a minimal bacterial genome. Science *351*, aad6253. https://doi.org/10.1126/science.aad6253.
15. Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., Vollger, M.R., Altemose, N., Uralsky, L., Gershman, A., et al. (2022). The complete sequence of a human genome. Science *376*, 44–53. https://doi.org/10.1126/science.abj6987.
16. Rocha, J.J., Jayaram, S.A., Stevens, T.J., Muschalik, N., Shah, R.D., Emran, S., Robles, C., Freeman, M., and Munro, S. (2023). Functional unknomics: systematic screening of conserved genes of unknown function. PLoS Biol. *21*, e3002222. https://doi.org/10.1371/journal.pbio.3002222.
17. Rodriguez, E.L., Poddar, S., Iftekhar, S., Suh, K., Woolfork, A.G., Ovbude, S., Pekarek, A., Walters, M., Lott, S., and Hage, D.S. (2020). Affinity chromatography: a review of trends and developments over the past 50 years. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. *1157*, 122332. https://doi.org/10.1016/j.jchromb.2020.122332.
18. Walters, R.R. (1985). Affinity chromatography. Anal. Chem. *57*, 1099A–1106A. https://doi.org/10.1021/ac00288a001.
19. Starkenstein, E. (1910). Ferment action and the influence upon it of neutral salts. Biochem. Z. *24*, 210–218.

20. D'Alessandro, G., and Sofia, F. (1935). The adsorption of antibodies from the sera of syphilitics and tuberculosis patients. Z Lmmunitats *84*, 237–250.

21. Meyer, K., and Pic, A. (1936). Isolation of antibodies by fixation on an adsorbent-antigen system with subsequent regeneration. Ann. Inst. Pasteur *56*, 401–412.22.

22. Campbell, D.H., Luescher, E., and Lerman, L.S. (1951). Immunologic Adsorbents: I. Isolation of Antibody by Means of a Cellulose-Protein Antigen. Proc. Natl. Acad. Sci. USA *37*, 575–578. https://doi.org/10.1073/pnas.37.9.575.

23. Lerman, L.S. (1953). Antibody chromatography on an immunologically specific adsorbent. Nature *172*, 635–636. https://doi.org/10.1038/172635a0.

24. Porath, J., and Axén, R. (1976). Immobilization of enzymes to agar, agarose, and Sephadex supports. Methods Enzymol. *44*, 19–45. https://doi.org/10.1016/s0076-6879(76)44005-3.

25. Porath, J., Axen, R., and Ernback, S. (1967). Chemical coupling of proteins to agarose. Nature *215*, 1491–1492. https://doi.org/10.1038/2151491a0.

26. Cuatrecasas, P., Wilchek, M., and Anfinsen, C.B. (1968). Selective enzyme purification by affinity chromatography. Proc. Natl. Acad. Sci. USA *61*, 636–643. https://doi.org/10.1073/pnas.61.2.636.

27. Formosa, T., Barry, J., Alberts, B.M., and Greenblatt, J. (1991). Using protein affinity chromatography to probe structure of protein machines. Methods Enzymol. *208*, 24–45. https://doi.org/10.1016/0076-6879(91)08005-3.

28. Greenblatt, J., and Li, J. (1981). The nusA gene protein of Escherichia coli. Its identification and a demonstration that it interacts with the gene N transcription anti-termination protein of bacteriophage lambda. J. Mol. Biol. *147*, 11–23. https://doi.org/10.1016/0022-2836(81)90076-0.

29. Roberts, J.W. (1969). Termination factor for RNA synthesis. Nature *224*, 1168–1174. https://doi.org/10.1038/2241168a0.

30. Luzzati, D. (1970). Regulation of lambda exonuclease synthesis: role of the N gene product and lambda repressor. J. Mol. Biol. *49*, 515–519. https://doi.org/10.1016/0022-2836(70)90261-5.

31. Greenblatt, J. (1972). Positive control of endolysin synthesis in vitro by the gene N protein of phage lambda. Proc. Natl. Acad. Sci. USA *69*, 3606–3610. https://doi.org/10.1073/pnas.69.12.3606.

32. Friedman, D.I., Baumann, M., and Baron, L.S. (1976). Cooperative effects of bacterial mutations affecting lambda N gene expression. I. Isolation and characterization of a nusB mutant. Virology *73*, 119–127. https://doi.org/10.1016/0042-6822(76)90066-0.

33. Georgopoulos, C.P., Swindle, J., Keppel, F., Ballivet, M., Bisig, R., and Eisen, H. (1980). Studies on the E. coli groNB (nusB) gene which affects bacteriophage lambda N gene function. Mol. Gen. Genet. *179*, 55–61. https://doi.org/10.1007/BF00268446.

34. Greenblatt, J., Malnoe, P., and Li, J. (1980). Purification of the gene N transcription anti-termination protein of bacteriophage lambda. J. Biol. Chem. *255*, 1465–1470. https://doi.org/10.1016/S0021-9258(19)86053-2.

35. Friedman, D.I., and Baron, L.S. (1974). Genetic characterization of a bacterial locus involved in the activity of the N function of phage lambda. Virology *58*, 141–148. https://doi.org/10.1016/0042-6822(74)90149-4.

36. Greenblatt, J., and Li, J. (1981). Interaction of the sigma factor and the nusA gene protein of E. coli with RNA polymerase in the initiation-termination cycle of transcription. Cell *24*, 421–428. https://doi.org/10.1016/0092-8674(81)90332-9.

37. Friedman, D.I., Schauer, A.T., Baumann, M.R., Baron, L.S., and Adhya, S.L. (1981). Evidence that ribosomal protein S10 participates in control of transcription termination. Proc. Natl. Acad. Sci. USA *78*, 1115–1118. https://doi.org/10.1073/pnas.78.2.1115.

38. Li, J., Horwitz, R., McCracken, S., and Greenblatt, J. (1992). NusG, a new Escherichia coli elongation factor involved in transcriptional antitermination by the N protein of phage lambda. J. Biol. Chem. *267*, 6012–6019. https://doi.org/10.1016/S0021-9258(18)42655-5.

39. Sullivan, S.L., and Gottesman, M.E. (1992). Requirement for E. coli NusG protein in factor-dependent transcription termination. Cell *68*, 989–994. https://doi.org/10.1016/0092-8674(92)90041-a.

40. Mason, S.W., Li, J., and Greenblatt, J. (1992). Direct interaction between two Escherichia coli transcription antitermination factors, NusB and ribosomal protein S10. J. Mol. Biol. *223*, 55–66. https://doi.org/10.1016/0022-2836(92)90715-v.

41. Li, J., Mason, S.W., and Greenblatt, J. (1993). Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. Genes Dev. *7*, 161–172. https://doi.org/10.1101/gad.7.1.161.

42. Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K., and Das, A. (1995). Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. Proc. Natl. Acad. Sci. USA *92*, 4061–4065. https://doi.org/10.1073/pnas.92.9.4061.

43. Mogridge, J., Mah, T.F., and Greenblatt, J. (1995). A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. Genes Dev. *9*, 2831–2845. https://doi.org/10.1101/gad.9.22.2831.

44. Legault, P., Li, J., Mogridge, J., Kay, L.E., and Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell *93*, 289–299. https://doi.org/10.1016/s0092-8674(00)81579-2.

45. Nodwell, J.R., and Greenblatt, J. (1993). Recognition of boxA antiterminator RNA by the E. coli antitermination factors NusB and ribosomal protein S10. Cell *72*, 261–268. https://doi.org/10.1016/0092-8674(93)90665-d.

46. Alberts, B.M., Barry, J., Bedinger, P., Formosa, T., Jongeneel, C.V., and Kreuzer, K.N. (1983). Studies on DNA replication in the bacteriophage T4 in vitro system. Cold Spring Harb. Symp. Quant. Biol. *47*, 655–668. https://doi.org/10.1101/sqb.1983.047.01.077.

47. Formosa, T., Burke, R.L., and Alberts, B.M. (1983). Affinity purification of bacteriophage T4 proteins essential for DNA replication and genetic recombination. Proc. Natl. Acad. Sci. USA *80*, 2442–2446. https://doi.org/10.1073/pnas.80.9.2442.

48. Formosa, T., and Alberts, B.M. (1984). The use of affinity chromatography to study proteins involved in bacteriophage T4 genetic recombination. Cold Spring Harb. Symp. Quant. Biol. *49*, 363–370. https://doi.org/10.1101/sqb.1984.049.01.043.

49. Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.-L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., et al. (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180–183. https://doi.org/10.1038/415180a.

50. Miller, K.G., Field, C.M., and Alberts, B.M. (1989). Actin-binding proteins from Drosophila embryos: a complex network of interacting proteins detected by F-actin affinity chromatography. J. Cell Biol. *109*, 2963–2975. https://doi.org/10.1083/jcb.109.6.2963.

51. Kellogg, D.R., Field, C.M., and Alberts, B.M. (1989). Identification of microtubule-associated proteins in the centrosome, spindle, and kinetochore of the early Drosophila embryo. J. Cell Biol. *109*, 2977–2991. https://doi.org/10.1083/jcb.109.6.2977.

52. Sopta, M., Carthew, R.W., and Greenblatt, J. (1985). Isolation of three proteins that bind to mammalian RNA polymerase II. J. Biol. Chem. *260*, 10353–10360. https://doi.org/10.1016/S0021-9258(17)39255-4.

53. Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G., and Greenblatt, J. (1988). RAP30/74: a general initiation factor that binds to RNA polymerase II. Mol. Cell. Biol. *8*, 1602–1613. https://doi.org/10.1128/mcb.8.4.1602-1613.1988.

54. Lane, D.P., and Crawford, L.V. (1979). T antigen is bound to a host protein in SV40-transformed cells. Nature *278*, 261–263. https://doi.org/10.1038/278261a0.

55. Linzer, D.I., and Levine, A.J. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. Cell *17*, 43–52. https://doi.org/10.1016/0092-8674(79)90293-9.

56. Whyte, P., Buchkovich, K.J., Horowitz, J.M., Friend, S.H., Raybuck, M., Weinberg, R.A., and Harlow, E. (1988). Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature *334*, 124–129. https://doi.org/10.1038/334124a0.

57. Reynolds, A.B., Kanner, S.B., Wang, H.C., and Parsons, J.T. (1989). Stable association of activated pp60src with two tyrosine-phosphorylated cellular proteins. Mol. Cell. Biol. *9*, 3951–3958. https://doi.org/10.1128/mcb.9.9.3951-3958.1989.

58. Koch, C.A., Anderson, D., Moran, M.F., Ellis, C., and Pawson, T. (1991). SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science *252*, 668–674. https://doi.org/10.1126/science.1708916.

59. Dynlacht, B.D., Hoey, T., and Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell *66*, 563–576. https://doi.org/10.1016/0092-8674(81)90019-2.

60. Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene *67*, 31–40. https://doi.org/10.1016/0378-1119(88)90005-4.

61. di Guan, C., Li, P., Riggs, P.D., and Inouye, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. Gene *67*, 21–30. https://doi.org/10.1016/0378-1119(88)90004-2.

62. Anderson, D., Koch, C.A., Grey, L., Ellis, C., Moran, M.F., and Pawson, T. (1990). Binding of SH2 domains of phospholipase C gamma 1, GAP, and Src to activated growth factor receptors. Science *250*, 979–982. https://doi.org/10.1126/science.2173144.

63. Moran, M.F., Koch, C.A., Anderson, D., Ellis, C., England, L., Martin, G.S., and Pawson, T. (1990). Src homology region 2 domains direct protein-protein interactions in signal transduction. Proc. Natl. Acad. Sci. USA *87*, 8622–8626. https://doi.org/10.1073/pnas.87.21.8622.

64. Edwards, A.M., Arrowsmith, C.H., Christendat, D., Dharamsi, A., Friesen, J.D., Greenblatt, J.F., and Vedadi, M. (2000). Protein production: feeding the crystallographers and NMR spectroscopists. Nat. Struct. Biol. *7*, 970–972. https://doi.org/10.1038/80751.

65. Stevens, R.C. (2000). Design of high-throughput methods of protein production for structural biology. Structure *8*, R177–R185. https://doi.org/10.1016/s0969-2126(00)00193-3.

66. Edman, P., Högfeldt, E., Sillén, L.G., and Kinell, P.-O. (1950). Method for determination of the amino acid sequence in peptides. Acta Chem. Scand. *4*, 283–293. https://doi.org/10.3891/acta.chem.scand.04-0283.

67. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. Science *246*, 64–71. https://doi.org/10.1126/science.2675315.

68. Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal. Chem. *60*, 2299–2301. https://doi.org/10.1021/ac00171a028.

69. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988). Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Comm. Mass Spectrometry *2*, 151–153. https://doi.org/10.1002/rcm.1290020802.

70. Mann, M., Højrup, P., and Roepstorff, P. (1993). Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol. Mass Spectrom. *22*, 338–345. https://doi.org/10.1002/bms.1200220605.

71. Yates, J.R., 3rd, Speicher, S., Griffin, P.R., and Hunkapiller, T. (1993). Peptide mass maps: a highly informative approach to protein identification. Anal. Biochem. *214*, 397–408. https://doi.org/10.1006/abio.1993.1514.

72. Dongré, A.R., Eng, J.K., and Yates, J.R., 3rd. (1997). Emerging tandem-mass-spectrometry techniques for the rapid identification of proteins. Trends Biotechnol. *15*, 418–425. https://doi.org/10.1016/S0167-7799(97)01110-4.

73. Richards, A.L., Eckhardt, M., and Krogan, N.J. (2021). Mass spectrometry-based protein-protein interaction networks for the study of human diseases. Mol. Syst. Biol. *17*, e8792. https://doi.org/10.15252/msb.20188792.

74. Howson, R., Huh, W.-K., Ghaemmaghami, S., Falvo, J.V., Bower, K., Belle, A., Dephoure, N., Wykoff, D.D., Weissman, J.S., and O'Shea, E.K. (2005). Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains. Comp. Funct. Genomics *6*, 2–16. https://doi.org/10.1002/cfg.449.

75. Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. USA *97*, 6640–6645. https://doi.org/10.1073/pnas.120163297.

76. Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000). An efficient recombination system for chromosome engineering in Escherichia coli. Proc. Natl. Acad. Sci. USA *97*, 5978–5983. https://doi.org/10.1073/pnas.100127597.

77. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat. Biotechnol. *17*, 1030–1032. https://doi.org/10.1038/13732.

78. Butland, G., Peregrín-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., et al. (2005). Interaction network containing conserved and essential protein complexes in Escherichia coli. Nature *433*, 531–537.

79. Gavin, A.-C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631–636. https://doi.org/10.1038/nature04532.

80. Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637–643. https://doi.org/10.1038/nature04670.

81. Hu, P., Janga, S.C., Babu, M., Díaz-Mejía, J.J., Butland, G., Yang, W., Pogoutse, O., Guo, X., Phanse, S., Wong, P., et al. (2009). Global functional atlas of Escherichia coli encompassing previously uncharacterized proteins. PLoS Biol. *7*, e96. https://doi.org/10.1371/journal.pbio.1000096.

82. Babu, M., Vlasblom, J., Pu, S., Guo, X., Graham, C., Bean, B.D.M., Burston, H.E., Vizeacoumar, F.J., Snider, J., Phanse, S., et al. (2012). Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. Nature *489*, 585–589. https://doi.org/10.1038/nature11354.

83. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. EMBO J. *19*, 1661–1671. https://doi.org/10.1093/emboj/19.7.1661.

84. Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc. Natl. Acad. Sci. USA *98*, 12902–12907. https://doi.org/10.1073/pnas.231473398.

85. Pijnappel, W.W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Séraphin, B., Aasland, R., and Stewart, A.F. (2001). The S. cerevisiae SET3 complex includes two histone

deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. Genes Dev. *15*, 2991–3004. https://doi.org/10.1101/gad.207401.

86. Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M., et al. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol. Cell *11*, 721–729. https://doi.org/10.1016/s1097-2765(03)00091-1.

87. Krogan, N.J., Keogh, M.-C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V., et al. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol. Cell *12*, 1565–1576. https://doi.org/10.1016/s1097-2765(03)00497-0.

88. Dziembowski, A., Ventura, A.-P., Rutz, B., Caspary, F., Faux, C., Halgand, F., Laprévote, O., and Séraphin, B. (2004). Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. EMBO J. *23*, 4847–4856. https://doi.org/10.1038/sj.emboj.7600482.

89. Krogan, N.J., Lam, M.H.Y., Fillingham, J., Keogh, M.-C., Gebbia, M., Li, J., Datta, N., Cagney, G., Buratowski, S., Emili, A., et al. (2004). Proteasome involvement in the repair of DNA double-strand breaks. Mol. Cell *16*, 1027–1034. https://doi.org/10.1016/j.molcel.2004.11.033.

90. Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517–522. https://doi.org/10.1038/nature03041.

91. Keogh, M.-C., Kim, J.-A., Downey, M., Fillingham, J., Chowdhury, D., Harrison, J.C., Onishi, M., Datta, N., Galicia, S., Emili, A., et al. (2006). A phosphatase complex that dephosphorylates gammaH2AX regulates DNA damage checkpoint recovery. Nature *439*, 497–501. https://doi.org/10.1038/nature04384.

92. Keogh, M.-C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593–605. https://doi.org/10.1016/j.cell.2005.10.025.

93. Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.-P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat. Methods *10*, 730–736. https://doi.org/10.1038/nmeth.2557.

94. Teo, G., Liu, G., Zhang, J., Nesvizhskii, A.I., Gingras, A.-C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. J. Proteomics *100*, 37–43. https://doi.org/10.1016/j.jprot.2013.10.023.

95. Sowa, M.E., Bennett, E.J., Gygi, S.P., and Harper, J.W. (2009). Defining the human deubiquitinating enzyme interaction landscape. Cell *138*, 389–403. https://doi.org/10.1016/j.cell.2009.04.042.

96. Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., McGovern, K.E., Clarke, S.C., Shales, M., Mercenne, G., Pache, L., Li, K., et al. (2011). Global landscape of HIV-human protein complexes. Nature *481*, 365–370. https://doi.org/10.1038/nature10719.

97. Guruharsha, K.G., Rual, J.-F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N., Beekman, C., Wong, C., Rhee, D.Y., Cenaj, O., et al. (2011). A protein complex network of Drosophila melanogaster. Cell *147*, 690–703. https://doi.org/10.1016/j.cell.2011.08.047.

98. Yang, X., Boehm, J.S., Yang, X., Salehi-Ashtiani, K., Hao, T., Shen, Y., Lubonja, R., Thomas, S.R., Alkan, O., Bhimdi, T., et al. (2011). A public genome-scale lentiviral expression library of human ORFs. Nat. Methods *8*, 659–661. https://doi.org/10.1038/nmeth.1638.

99. Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., et al. (2015). The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell *162*, 425–440. https://doi.org/10.1016/j.cell.2015.06.043.

100. Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F., et al. (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *163*, 712–723. https://doi.org/10.1016/j.cell.2015.09.053.

101. Li, T., Wernersson, R., Hansen, R.B., Horn, H., Mercer, J., Slodkowicz, G., Workman, C.T., Rigina, O., Rapacki, K., Stærfeldt, H.H., et al. (2017). A scored human protein-protein interaction network to catalyze genomic interpretation. Nat. Methods *14*, 61–64. https://doi.org/10.1038/nmeth.4083.

102. Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Ruepp, A. (2019). CORUM: the comprehensive resource of mammalian protein complexes-2019. Nucleic Acids Res. *47*, D559–D563. https://doi.org/10.1093/nar/gky973.

103. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.-J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F., et al. (2021). The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci. *30*, 187–200. https://doi.org/10.1002/pro.3978.

104. Buljan, M., Ciuffa, R., van Drogen, A., Vichalkovski, A., Mehnert, M., Rosenberger, G., Lee, S., Varjosalo, M., Pernas, L.E., Spegg, V., et al. (2020). Kinase Interaction Network Expands Functional and Disease Roles of Human Kinases. Mol. Cell *79*, 504–520.e9. https://doi.org/10.1016/j.molcel.2020.07.001.

105. Couzens, A.L., Knight, J.D.R., Kean, M.J., Teo, G., Weiss, A., Dunham, W.H., Lin, Z.-Y., Bagshaw, R.D., Sicheri, F., Pawson, T., et al. (2013). Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. Sci. Signal. *6*, rs15. https://doi.org/10.1126/scisignal.2004712.

106. Ma, S., Meng, Z., Chen, R., and Guan, K.-L. (2019). The Hippo Pathway: Biology and Pathophysiology. Annu. Rev. Biochem. *88*, 577–604. https://doi.org/10.1146/annurev-biochem-013118-111829.

107. Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.-C., Hirai, A., et al. (2006). Large-scale identification of protein-protein interaction of Escherichia coli K-12. Genome Res. *16*, 686–691. https://doi.org/10.1101/gr.4527806.

108. Cho, N.H., Cheveralls, K.C., Brunner, A.-D., Kim, K., Michaelis, A.C., Raghavan, P., Kobayashi, H., Savy, L., Li, J.Y., Canaj, H., et al. (2022). OpenCell: Endogenous tagging for the cartography of human cellular organization. Science *375*, eabi6983. https://doi.org/10.1126/science.abi6983.

109. Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A census of human soluble protein complexes. Cell *150*, 1068–1081. https://doi.org/10.1016/j.cell.2012.08.011.

110. Kristensen, A.R., Gsponer, J., and Foster, L.J. (2012). A high-throughput approach for measuring temporal changes in the interactome. Nat. Methods *9*, 907–909. https://doi.org/10.1038/nmeth.2131.

111. Pourhaghighi, R., Ash, P.E.A., Phanse, S., Goebels, F., Hu, L.Z.M., Chen, S., Zhang, Y., Wierbowski, S.D., Boudeau, S., Moutaoufik, M.T., et al. (2020). BrainMap elucidates the macromolecular connectivity landscape of mammalian brain. Cell Syst. *10*, 333–350.e14.

112. Fossati, A., Mozumdar, D., Kokontis, C., Mèndez-Moran, M., Nieweglowska, E., Pelin, A., Li, Y., Guo, B., Krogan, N.J., Agard, D.A., et al. (2023). Next-generation proteomics for quantitative Jumbophage-bacteria interaction mapping. Nat. Commun. *14*, 5156. https://doi.org/10.1038/s41467-023-40724-w.

113. Havugimana, P.C., Goel, R.K., Phanse, S., Youssef, A., Padhorny, D., Kotelnikov, S., Kozakov, D., and Emili, A. (2022). Scalable multiplex co-fractionation/mass spectrometry platform for accelerated protein interactome discovery. Nat. Commun. *13*, 4043. https://doi.org/10.1038/s41467-022-31809-z.

114. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression

proteomics. Mol. Cell. Proteomics *1*, 376–386. https://doi.org/10.1074/mcp.m200025-mcp200.

115. Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B.C., and Aebersold, R. (2018). Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. Mol. Syst. Biol. *14*, e8126. https://doi.org/10.15252/msb.20178126.

116. Li, J., Cai, Z., Bomgarden, R.D., Pike, I., Kuhn, K., Rogers, J.C., Roberts, T.M., Gygi, S.P., and Paulo, J.A. (2021). TMTpro-18plex: The Expanded and Complete Set of TMTpro Reagents for Sample Multiplexing. J. Proteome Res. *20*, 2964–2972. https://doi.org/10.1021/acs.jproteome.1c00168.

117. Mateus, A., Hevler, J., Bobonis, J., Kurzawa, N., Shah, M., Mitosch, K., Goemans, C.V., Helm, D., Stein, F., Typas, A., et al. (2020). The functional proteome landscape of Escherichia coli. Nature *588*, 473–478. https://doi.org/10.1038/s41586-020-3002-5.

118. Förster, T. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz. Annalen der Physik *437*, 55–75. https://doi.org/10.1002/andp.19484370105.

119. Pfleger, K.D.G., and Eidne, K.A. (2006). Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). Nat. Methods *3*, 165–174. https://doi.org/10.1038/nmeth841.

120. Miyawaki, A., Llopis, J., Heim, R., McCaffery, J.M., Adams, J.A., Ikura, M., and Tsien, R.Y. (1997). Fluorescent indicators for Ca2+based on green fluorescent proteins and calmodulin. Nature *388*, 882–887. https://doi.org/10.1038/42264.

121. Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gústafsdóttir, S.M., Östman, A., and Landegren, U. (2002). Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol. *20*, 473–477. https://doi.org/10.1038/nbt0502-473.

122. Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.-J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.-G., et al. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nat. Methods *3*, 995–1000. https://doi.org/10.1038/nmeth947.

123. Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. *196*, 801–810. https://doi.org/10.1083/jcb.201112098.

124. Rhee, H.-W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science *339*, 1328–1331. https://doi.org/10.1126/science.1230593.

125. Branon, T.C., Bosch, J.A., Sanchez, A.D., Udeshi, N.D., Svinkina, T., Carr, S.A., Feldman, J.L., Perrimon, N., and Ting, A.Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. Nat. Biotechnol. *36*, 880–887. https://doi.org/10.1038/nbt.4201.

126. Go, C.D., Knight, J.D.R., Rajasekharan, A., Rathod, B., Hesketh, G.G., Abe, K.T., Youn, J.-Y., Samavarchi-Tehrani, P., Zhang, H., Zhu, L.Y., et al. (2021). A proximity-dependent biotinylation map of a human cell. Nature *595*, 120–124. https://doi.org/10.1038/s41586-021-03592-2.

127. Seath, C.P., Burton, A.J., Sun, X., Lee, G., Kleiner, R.E., MacMillan, D.W.C., and Muir, T.W. (2023). Tracking chromatin state changes using nanoscale photo-proximity labelling. Nature *616*, 574–580. https://doi.org/10.1038/s41586-023-05914-y.

128. Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245–246. https://doi.org/10.1038/340245a0.

129. Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., et al. (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature *403*, 623–627. https://doi.org/10.1038/35001009.

130. Walhout, A.J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods *24*, 297–306. https://doi.org/10.1006/meth.2001.1190.

131. Luck, K., Kim, D.-K., Lambourne, L., Spirohn, K., Begg, B.E., Bian, W., Brignall, R., Cafarelli, T., Campos-Laborie, F.J., Charloteaux, B., et al. (2020). A reference map of the human binary protein interactome. Nature *580*, 402–408. https://doi.org/10.1038/s41586-020-2188-x.

132. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242. https://doi.org/10.1093/nar/28.1.235.

133. Johnsson, N., and Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions in vivo. Proc. Natl. Acad. Sci. USA *91*, 10340–10344. https://doi.org/10.1073/pnas.91.22.10340.

134. Shekhawat, S.S., and Ghosh, I. (2011). Split-protein systems: beyond binary protein-protein interactions. Curr. Opin. Chem. Biol. *15*, 789–797. https://doi.org/10.1016/j.cbpa.2011.10.014.

135. Stagljar, I., Korostensky, C., Johnsson, N., and te Heesen, S. (1998). A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. Proc. Natl. Acad. Sci. USA *95*, 5187–5192. https://doi.org/10.1073/pnas.95.9.5187.

136. Rappsilber, J., Siniossoglou, S., Hurt, E.C., and Mann, M. (2000). A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. Anal. Chem. *72*, 267–275. https://doi.org/10.1021/ac991081o.

137. Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2016). Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. Trends Biochem. Sci. *41*, 20–32. https://doi.org/10.1016/j.tibs.2015.10.008.

138. Kühlbrandt, W. (2014). Biochemistry. The resolution revolution. Science *343*, 1443–1444. https://doi.org/10.1126/science.1251652.

139. Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F.-U., Ban, N., et al. (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. Science *337*, 1348–1352. https://doi.org/10.1126/science.1221483.

140. Ideker, T., and Krogan, N.J. (2012). Differential network biology. Mol. Syst. Biol. *8*, 565. https://doi.org/10.1038/msb.2011.99.

141. Lobingier, B.T., Hüttenhain, R., Eichel, K., Miller, K.B., Ting, A.Y., von Zastrow, M., and Krogan, N.J. (2017). An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. Cell *169*, 350–360.e12. https://doi.org/10.1016/j.cell.2017.03.022.

142. Lambert, J.-P., Ivosev, G., Couzens, A.L., Larsen, B., Taipale, M., Lin, Z.-Y., Zhong, Q., Lindquist, S., Vidal, M., Aebersold, R., et al. (2013). Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. Nat. Methods *10*, 1239–1245. https://doi.org/10.1038/nmeth.2702.

143. Kratz, A., Kim, M., Kelly, M.R., Zheng, F., Koczor, C.A., Li, J., Ono, K., Qin, Y., Churas, C., Chen, J., et al. (2023). A multi-scale map of protein assemblies in the DNA damage response. Cell Syst. *14*, 447–463.e8. https://doi.org/10.1016/j.cels.2023.04.007.

144. Hota, S.K., Johnson, J.R., Verschueren, E., Thomas, R., Blotnick, A.M., Zhu, Y., Sun, X., Pennacchio, L.A., Krogan, N.J., and Bruneau, B.G. (2019). Dynamic BAF chromatin remodeling complex subunit inclusion promotes temporally distinct gene expression programs in cardiogenesis. Development *146*, dev174086. https://doi.org/10.1242/dev.174086.

145. Hosp, F., Vossfeldt, H., Heinig, M., Vasiljevic, D., Arumughan, A., Wyler, E., Genetic and Environmental Risk for Alzheimer's Disease GERAD1 Consortium, Landthaler, M., Hubner, N., Wanker, E.E., et al. (2015). Quantitative interaction proteomics of neurodegenerative disease proteins. Cell Rep *11*, 1134–1146.

146. Kelly, M.R., Kostyrko, K., Han, K., Mooney, N.A., Jeng, E.E., Spees, K., Dinh, P.T., Abbott, K.L., Gwinn, D.M., Sweet-Cordero, E.A., et al. (2020). Combined Proteomic and Genetic Interaction Mapping Reveals

New RAS Effector Pathways and Susceptibilities. Cancer Discov. *10*, 1950–1967. https://doi.org/10.1158/2159-8290.CD-19-1274.

147. O'Connor, C.M., Leonard, D., Wiredja, D., Avelar, R.A., Wang, Z., Schlatzer, D., Bryson, B., Tokala, E., Taylor, S.E., Upadhyay, A., et al. (2020). Inactivation of PP2A by a recurrent mutation drives resistance to MEK inhibitors. Oncogene *39*, 703–717. https://doi.org/10.1038/s41388-019-1012-2.

148. Swaney, D.L., Ramms, D.J., Wang, Z., Park, J., Goto, Y., Soucheray, M., Bhola, N., Kim, K., Zheng, F., Zeng, Y., et al. (2021). A protein network map of head and neck cancer reveals PIK3CA mutant drug sensitivity. Science *374*, eabf2911. https://doi.org/10.1126/science.abf2911.

149. Zheng, F., Kelly, M.R., Ramms, D.J., Heintschel, M.L., Tao, K., Tutuncuoglu, B., Lee, J.J., Ono, K., Foussard, H., Chen, M., et al. (2021). Interpretation of cancer mutations using a multiscale map of protein systems. Science *374*, eabf3067. https://doi.org/10.1126/science.abf3067.

150. Kim, M., Park, J., Bouhaddou, M., Kim, K., Rojc, A., Modak, M., Soucheray, M., McGregor, M.J., O'Leary, P., Wolf, D., et al. (2021). A protein interaction landscape of breast cancer. Science *374*, eabf3066. https://doi.org/10.1126/science.abf3066.

151. Wang, B., Vartak, R., Zaltsman, Y., Naing, Z.Z.C., Hennick, K.M., Polacco, B.J., Bashir, A., Eckhardt, M., Bouhaddou, M., Xu, J., et al. (2024). A foundational atlas of autism protein interactions reveals molecular convergence. Preprint at bioRxiv. https://doi.org/10.1101/2023.12.03.569805.

152. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature *583*, 459–468. https://doi.org/10.1038/s41586-020-2286-9.

153. Bouhaddou, M., Reuschl, A.-K., Polacco, B.J., Thorne, L.G., Ummadi, M.R., Ye, C., Rosales, R., Pelin, A., Batra, J., Jang, G.M., et al. (2023). SARS-CoV-2 variants evolve convergent strategies to remodel the host response. Cell *186*, 4597–4614.e26. https://doi.org/10.1016/j.cell.2023.08.026.

154. Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M.P., et al. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592. https://doi.org/10.1016/j.cell.2005.10.023.

155. Hodges, C., Kirkland, J.G., and Crabtree, G.R. (2016). The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. Cold Spring Harb. Perspect. Med. *6*, a026930. https://doi.org/10.1101/cshperspect.a026930.

156. Malone, H.A., and Roberts, C.W.M. (2024). Chromatin remodellers as therapeutic targets. Nat. Rev. Drug Discov. *23*, 661–681. https://doi.org/10.1038/s41573-024-00978-5.

157. Valencia, A.M., Sankar, A., van der Sluijs, P.J., Satterstrom, F.K., Fu, J., Talkowski, M.E., Vergano, S.A.S., Santen, G.W.E., and Kadoch, C. (2023). Landscape of mSWI/SNF chromatin remodeling complex perturbations in neurodevelopmental disorders. Nat. Genet. *55*, 1400–1412. https://doi.org/10.1038/s41588-023-01451-6.

158. He, S., Wu, Z., Tian, Y., Yu, Z., Yu, J., Wang, X., Li, J., Liu, B., and Xu, Y. (2020). Structure of nucleosome-bound human BAF complex. Science *367*, 875–881. https://doi.org/10.1126/science.aaz9761.

159. Noddings, C.M., Johnson, J.L., and Agard, D.A. (2023). Cryo-EM reveals how Hsp90 and FKBP immunophilins co-regulate the glucocorticoid receptor. Nat. Struct. Mol. Biol. *30*, 1867–1877. https://doi.org/10.1038/s41594-023-01128-y.

160. Wang, F., Liu, Y., Yu, Z., Li, S., Feng, S., Cheng, Y., and Agard, D.A. (2020). General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. Proc. Natl. Acad. Sci. USA *117*, 24269–24273. https://doi.org/10.1073/pnas.2009707117.

161. Levy, A., Grzadkowski, M., Poitevin, F., Vallese, F., Clarke, O.B., Wetzstein, G., and Zhong, E.D. (2024). Revealing biomolecular structure and motion with neural ab initio cryo-EM reconstruction. Preprint at bioRxiv. https://doi.org/10.1101/2024.05.30.596729.

162. Asano, S., Fukuda, Y., Beck, F., Aufderheide, A., Förster, F., Danev, R., and Baumeister, W. (2015). Proteasomes. A molecular census of 26S proteasomes in intact neurons. Science *347*, 439–442. https://doi.org/10.1126/science.1261197.

163. Young, L.N., and Villa, E. (2023). Bringing Structure to Cell Biology with Cryo-Electron Tomography. Annu. Rev. Biophys. *52*, 573–595. https://doi.org/10.1146/annurev-biophys-111622-091327.

164. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

165. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature *596*, 590–596. https://doi.org/10.1038/s41586-021-03828-1.

166. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

167. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2022). Protein complex prediction with AlphaFold-Multimer. Preprint at bioRxiv. https://doi.org/10.1101/2021.10.04.463034.

168. Humphreys, I.R., Pei, J., Baek, M., Krishnakumar, A., Anishchenko, I., Ovchinnikov, S., Zhang, J., Ness, T.J., Banjade, S., Bagde, S.R., et al. (2021). Computed structures of core eukaryotic protein complexes. Science *374*, eabm4805. https://doi.org/10.1126/science.abm4805.

169. Burke, D.F., Bryant, P., Barrio-Hernandez, I., Memon, D., Pozzati, G., Shenoy, A., Zhu, W., Dunham, A.S., Albanese, P., Keller, A., et al. (2023). Towards a structurally resolved human protein interaction network. Nat. Struct. Mol. Biol. *30*, 216–225. https://doi.org/10.1038/s41594-022-00910-8.

170. Lim, Y., Tamayo-Orrego, L., Schmid, E., Tarnauskaite, Z., Kochenova, O.V., Gruar, R., Muramatsu, S., Lynch, L., Schlie, A.V., Carroll, P.L., et al. (2023). In silico protein interaction screening uncovers DONSON's role in replication initiation. Science *381*, eadi3448. https://doi.org/10.1126/science.adi3448.

171. Tsai, J.M., Nowak, R.P., Ebert, B.L., and Fischer, E.S. (2024). Targeted protein degradation: from mechanisms to clinic. Nat. Rev. Mol. Cell Biol. *25*, 740–757. https://doi.org/10.1038/s41580-024-00729-9.

172. Truong, T., and Kelly, R.T. (2024). What's new in single-cell proteomics. Curr. Opin. Biotechnol. *86*, 103077. https://doi.org/10.1016/j.copbio.2024.103077.

173. Zhu, Y., Akkaya, K.C., Ruta, J., Yokoyama, N., Wang, C., Ruwolt, M., Lima, D.B., Lehmann, M., and Liu, F. (2024). Cross-link assisted spatial proteomics to map sub-organelle proteomes and membrane protein topologies. Nat. Commun. *15*, 3290. https://doi.org/10.1038/s41467-024-47569-x.

174. Kaake, R.M., Wang, X., Burke, A., Yu, C., Kandur, W., Yang, Y., Novtisky, E.J., Second, T., Duan, J., Kao, A., et al. (2014). A new in vivo cross-linking mass spectrometry platform to define protein-protein interactions in living cells. Mol. Cell. Proteomics *13*, 3533–3543. https://doi.org/10.1074/mcp.M114.042630.

175. Jiang, P.-L., Wang, C., Diehl, A., Viner, R., Etienne, C., Nandhikonda, P., Foster, L., Bomgarden, R.D., and Liu, F. (2022). A Membrane-Permeable and Immobilized Metal Affinity Chromatography (IMAC) Enrichable Cross-Linking Reagent to Advance In Vivo Cross-Linking Mass Spectrometry. Angew. Chem. Int. Ed. Engl. 61, e202113937.