Current Clinical and Medical Education

Received 30 Jun 2024 | Revised 29 Jul 2024 | Accepted 20 Aug 2024 | Published Online 6 Sep 2024

Published By: Vision Publisher

CCME 02 (9), 36-52

ISSN 2942-0792

Mapping of Protein–Protein Interaction: Identification, Affinity Purification, Tandem Affinity Purification and Quality of Protein Interaction Data

Zainab Mohammed Ali Abdul-Kadhim Shaker[1], Ghafran Qais Jawad Jarallah[2], Mohammed Hassan Kalf Saad[3], Safaa Ibrahim Kazim Jadou[4]

[1]University of Babylon, College of Science for Women, Department of Biology

[2]University of Babylon, College of Science / Department of Biology

[3]University of Al-Qadisiyah, College of Science, Department of Biology

[4]Rashid University College, Department of Biology

Abstract:
The utilization of mass spectrometry for protein identification has brought about a sea change in the proteomics discipline. Mass spectrometry, in conjunction with a number of affinity purification methods, can detect interactions between proteins. Protein complexes can be purified via tandem affinity purification or any of a number of other tags. Another popular tag is the FLAG tag; it is tiny and usually doesn't get in the way of the protein's activity. To prepare proteins for further identification using ESI-MS or MALDI-MS, various affinity purification techniques are employed. Rapid progress in the creation of new treatment strategies depends on our ability to better comprehend the biological pathways underpinning disease. Protein interactions are a common mediator of disease processes. We can learn more about the causes, development, and pathophysiology of diseases and find possible druggable targets if we can deduce how protein-protein interactions physically change in reaction to mutations, pathological circumstances, or pathogen infection. Recent developments in quantitative mass spectrometry (MS)-based methods have made it possible to map these alterations in protein-protein interactions produced by diseases on a worldwide scale in an unbiased manner. In this article, we take a look back at magnetic resonance imaging (MS) methods that have helped pinpoint system-level protein-protein interactions, and we talk about the problems with these approaches, as well as new developments in MS that try to fix them. Nevertheless, disease networks should not be considered independently. The importance of the mechanisms suggested by an interactome must be assessed, however, in the same way as with any systems biology approach. Because of its scalability, immortalised cell lines are a common tool for PPI research. Though simple to work with, these cell lines may miss the mark when it comes to capturing relationships that matter in more complicated tissues and creatures. More functionally relevant and physiologically correct disease models for studying interactions can be developed with the use of newer genetic techniques, such as CRISPR/Cas9-based genome engineering of primary cells. We are getting closer to incorporating these technologies into personalised medical applications as technology keeps getting better and these methods become more widely available and have higher throughput. In addition to helping doctors understand how the body works, they may also pinpoint exactly where a patient's network is most vulnerable and advise them on the best courses of treatment. Proteomics has progressed greatly since MS was used for protein identification. This area will see further advancements with the introduction of novel affinity purification methods and MS machines. The rate of false positives and negatives can be reduced through more stringent experimental design and data processing. At long last, everything is in place to go on with the human interactome identification. On a grand scale, scientists will one day be able to examine how protein interactions change in response to various stimuli and compare the interactome of cells in various disease states or after treatment with various cues.

Keywords: Protein–Protein Interaction, Affinity Purification, Affinity, Quality, Protein Interaction Data.

Corresponding Author: Zainab Mohammed Ali Abdul-Kadhim Shaker, University of Babylon, College of Science for Women, Department of Biology

Introduction

In the last ten years, data generated by various branches of the life sciences has skyrocketed. The completion, in 2003, of the human genome project was one such development. The human genome project shifted its emphasis from DNA to proteins, creating a new challenge: figuring out what each gene does. A thorough comprehension of the activity of proteins is crucial to comprehending cellular function since they make up the majority of the functional units in the cell. Understanding proteins in all their complexity, from their structures and changes to their localisation and interactions with other proteins, is the goal of proteomics [1]. Expression proteomics and functional proteomics are two subfields of proteomics. When compared to normal cells, expression proteomics examines how protein expression changes under various situations. Protein expression studies can compare normal cells to those subjected to various medications, stressors, and disease states. The use of two-dimensional gels has allowed this area of research to progress beyond the one-protein scale used in Western blot analysis. go on to the last step, which involves employing methods like isotope labelling to conduct large-scale analyses of changes in protein expression. Similar to SILAC, which stands for stable isotope labelling with amino acids in cell culture. Understanding protein activities and clarifying their cellular role is the focus of functional proteomics, the second main field in proteomics. The discovery of protein-protein interactions is now considered a crucial step in deciphering the roles played by various proteins. Reason being, the majority of proteins are not "island" proteins; rather, they perform their functions by building specific complexes in response to environmental cues. The cell's protein-protein interactions provide a key to unlocking the mysteries of many proteins whose roles remain a mystery. The application of mass spectrometry (MS) for protein identification is a significant development that has radically altered the proteomics landscape. Because biomolecules are big and polar ions, it is exceedingly difficult to move them to the gas phase, which has rendered the use of MS in biomolecule analysis very inefficient. In order to address these issues, the use of MS in biomolecule analysis had to be postponed until ionisation methods were developed [2]. The development of electrospray ionisation by John Bennett Fenn and matrix-assisted laser desorption/ionization (MALDI) by Koichi Tanaka in the last twenty years has made it possible to easily analyse these big molecules. Second, high-throughput bioinformatics for mass spectrometry-based protein and peptide analysis was born out of the abundance of sequences databases made possible by genome sequencing and annotation initiatives. Suites of separation methods and data analysis have also been developed, which further supports the usefulness of these approaches. The growing interest in multiple sclerosis and its development were greatly aided by these innovations taken as a whole. The next big thing in proteomic research is MS 69, which will help identify protein-protein interactions.

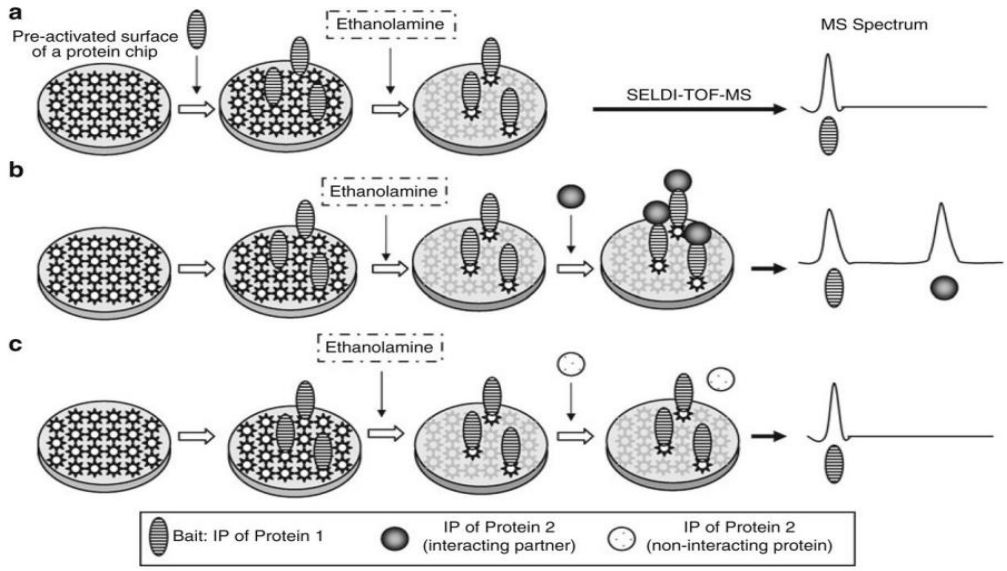

Figure 1. Analysis of protein-protein interactions using SELDI-TOF-MS on a ProteinChip array protected with ethanolamine is depicted schematically. (a) When only one protein or peptide is attached to the surface of the array, only one peak will be seen in the mass spectra. (b) Upon precise interaction between the bait and the

target protein or peptide, two peaks will be observed. (c) the second peak not appearing in the mass spec data because the subsequently applied protein or peptide did not interact with the bait protein or peptide.

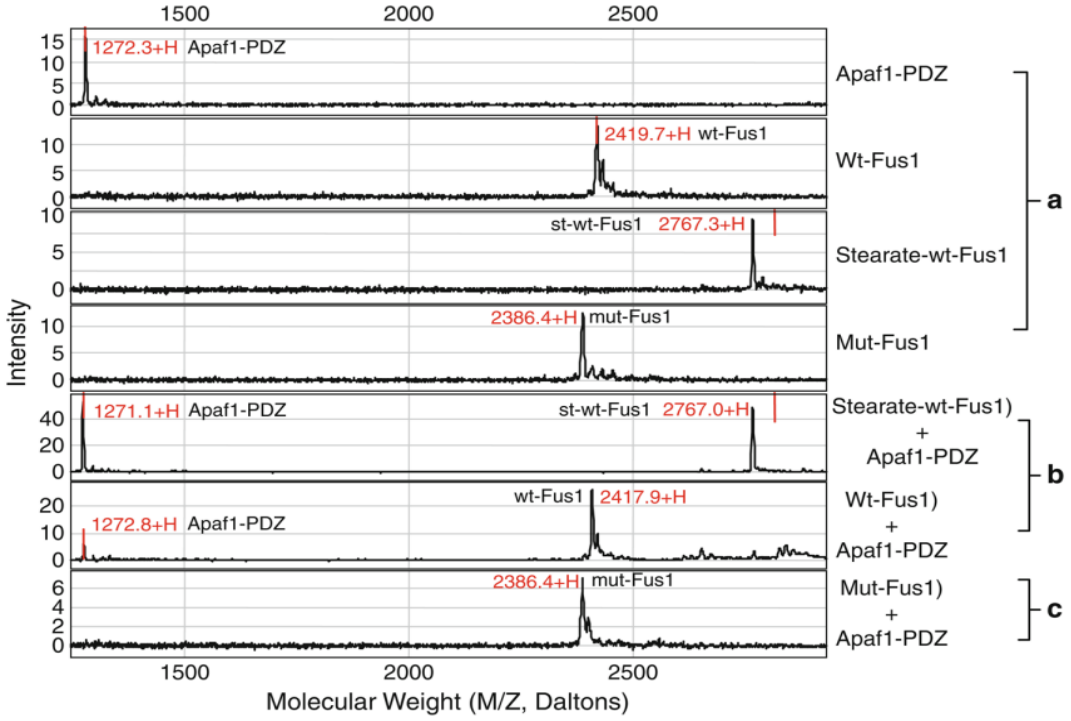

Figure 2. Using an ethanolamine protection test on a ProteinChip array with SELDI-MS, we examined the Fus1-Apaf-1 interaction with synthetic peptides that were created from predicted protein interaction motifs. Precision in mass measurements and comparisons with negative mutant and nonspecific control peptides allowed for the detection of the Fus1 PDZ domain's unique interaction with the Apaf-1 c-terminal peptide. Panels A show individual peaks of different Fus1 or Apaf1 peptides when loaded onto proteinchip arrays; Panels B show two distinct peaks only when the peptides interact specifically with each other; and Panels C show that when a mutation is introduced in the PDZ binding motif of Fus1, no Apaf1 peptide peak appears in the spectrum..

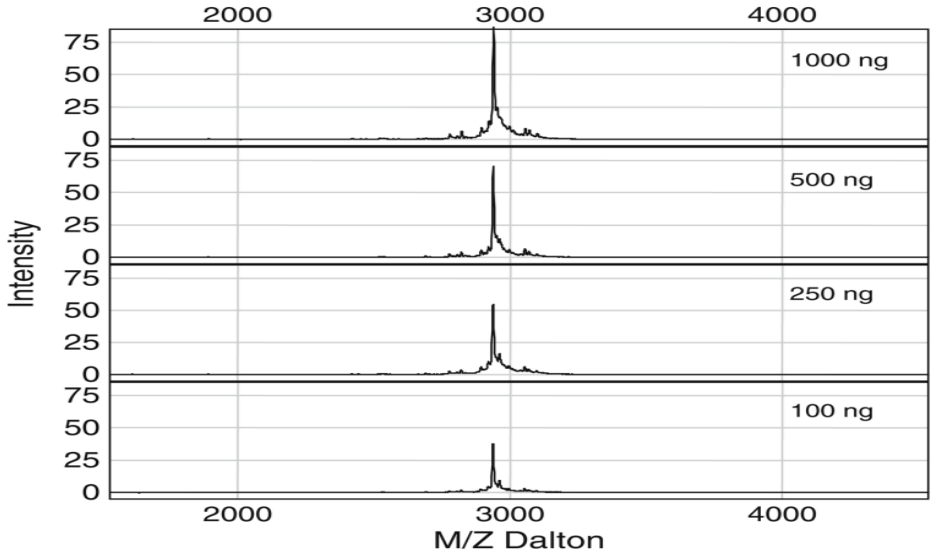

Figure 3. Analysing the PS10 ProteinChip array using SELDI-TOF-MS for peptide binding quantification and sensitivity determination. Each PS10 chip is loaded with a serial dilution of a known concentration of purified synthetic peptides, and the peak intensity is used to quantify the amount of peptide.

Exploring the Interaction Between Proteins

Although no-coding RNA is important as regulatory and functional elements, the human genome's smaller-than-expected number of genes has strengthened the hypothesis that proteins can have multifunction in the cell [4]. Depending on its localisation and the proteins it interacts with, a single protein's function can change. Moonlighting of proteins is an intriguing example that demonstrates the multifunctional characteristics. There are a plethora of methods for studying cellular protein-protein interactions. Two of the most popular high-throughput methods are affinity purification coupled to MS and yeast two hybrids (Y2H). Two domains found on transcription factors—an activation domain and a DNA binding domain—form the basis of Y2H tests. The two parts are isolated and then linked to two proteins that may interact with each other in this test. When these proteins bind, they activate a detectable reporter gene. When looking for evidence of protein-protein interaction, this method is frequently employed. A false-positive rate of up to 50% is the primary argument against Y2H [5]. The fact that the test looks at how over-expressed fusion proteins interact in the yeast nucleus is one possible explanation for the high false-positive rate. As an alternative, affinity purification linked to MS has seen extensive use. Here, an anti-tag system that has been immobilised on a solid support is utilised to purify protein complexes after tagging them with specific tags. The protein sample is subsequently digested into tiny peptides that can be detected by MS after being separated using one of many procedures. One benefit of this approach is the use of commercially accessible, very specific anti-tag devices in a variety of formats. The development of more reliable and less complicated molecular biology techniques has further simplified and facilitated this method. [6].

Protein Complex Affinity Purification

To find protein-protein interactions, several affinity-based protein purification techniques have been applied recently.

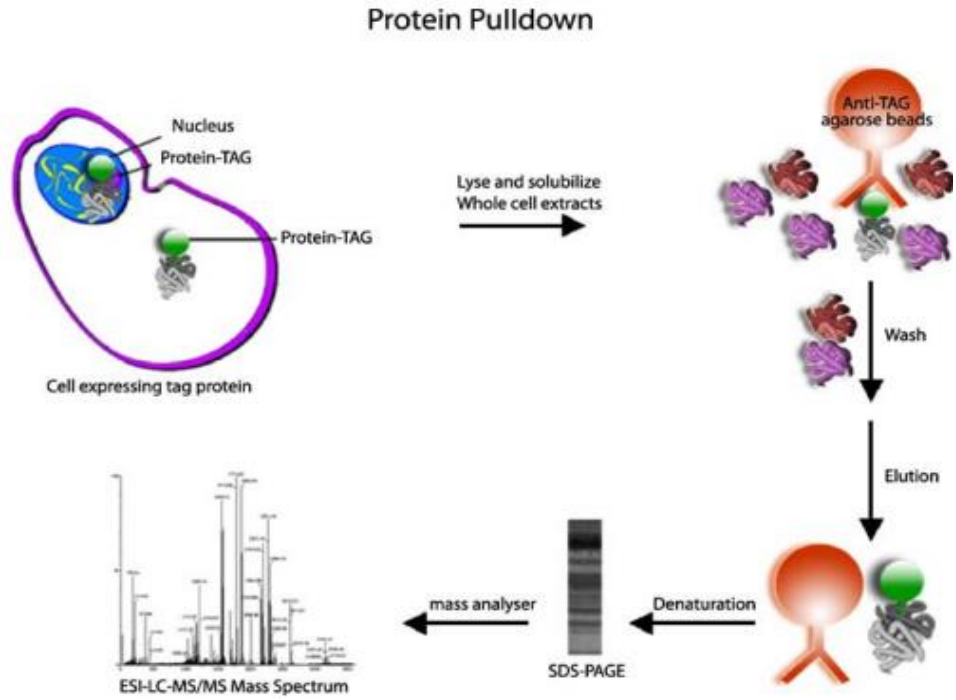

Figure 4. The expression of an affinity tag determines the several processes in the anti-tag protein purification method, as shown in the diagram. Short hydrophilic peptides, like FLAG, hemagglutinin (HA), or poly-His, make up these tags. Additional tags can be tiny proteins such as GST, thioredoxin, or GFP.

Purification of Tandem Affinity

The purpose of tandem affinity purification (TAP) was to create a way to isolate protein complexes that were naturally produced at physiological quantities [7]. As its name suggests, this technique depends on utilising two tags. The FLAG tag, two IgG-binding units of Staphylococcus aureus protein A (ProtA), the Strep tag, the His-tag, the calmodulin-binding peptide (CBP), and the chitin-binding domain (CBD) were among the tags that were examined in the initial study by Rigaut et al. (1999). Regardless, the protein function was unaffected by any of the tags. A protease recognition site for the tobacco etch virus (TEV) separates the two tags. The interaction partners of 589 proteins were identified by Gavin et al. (2002) using the TAP technique coupled to MS. During this investigation, 232 multi-protein complexes were discovered. In two separate investigations that examined all 6466 ORFs in yeast, the remaining yeast interactome was identified utilising TAP purification procedures. The proteins of interest were fused to TAP tags using homologous recombination in several investigations. As a result, these proteins can be expressed at physiological levels of tagged protein production, guided by their endogenous promoters. The tagged protein attaches to the IgG-sepharose via its ProtA tag when cellular lysates containing the protein are added to the IgG-sepharose [8]. It was necessary to wash the tagged protein and its binding partners in order to lower the level of contamination. In order to enhance the level of purity, the immobilised protein complex was exposed to TEV protease with the purpose of releasing the target protein and its associated interactors. The two-step purification process continues with the use of calmodulin-sepharose, which, when activated with calcium, binds to the CBP tag on the target protein. Following the washing step, EGTA was used to elute the protein complex. Alternate approaches can be employed for the resolution of eluted protein complexes. A one-dimensional SDS-PAGE gel was used in the study by Gavin et al. [9]. (2002). Next, the gel was dyed. After that, the bands of interest were proteolytically digested and MALDI-MS analysis was performed. In MALDI, the analyte sample and a matrix are combined in a high ratio to promote ionisation and the production of gas-phase molecules. Small organic molecules often make up the matrix. A metal substrate is spotted with a mixture of the matrix and the analyte at a high ratio to produce protonated gas-phase molecules. A laser beam is used to ionise the matrix after the crystals have dried. The theory is that the matrix transfers some of its charge to the analyte, causing it to become ionised. Simultaneously, the matrix shields the analyte from the laser's disruptive energy. While doubly charged ions are less common, MALDI ionisation can produce ions with a single charge. provides an outline of the process of identifying proteins using MALDI ionisation in conjunction with affinity purification. When working with yeast as a model system, there are several benefits, including as the ability to tag the target protein and have its expression controlled by its own promoter through homologous recombination. This method has the dual benefit of maintaining the protein's normal expression level and removing untagged proteins from the cell. To get around the lack of homologous recombination, various strategies have been employed in mammalian systems. Methods such as transient transfection, stable cell lines, and inducible promoters are examples of these. [9].

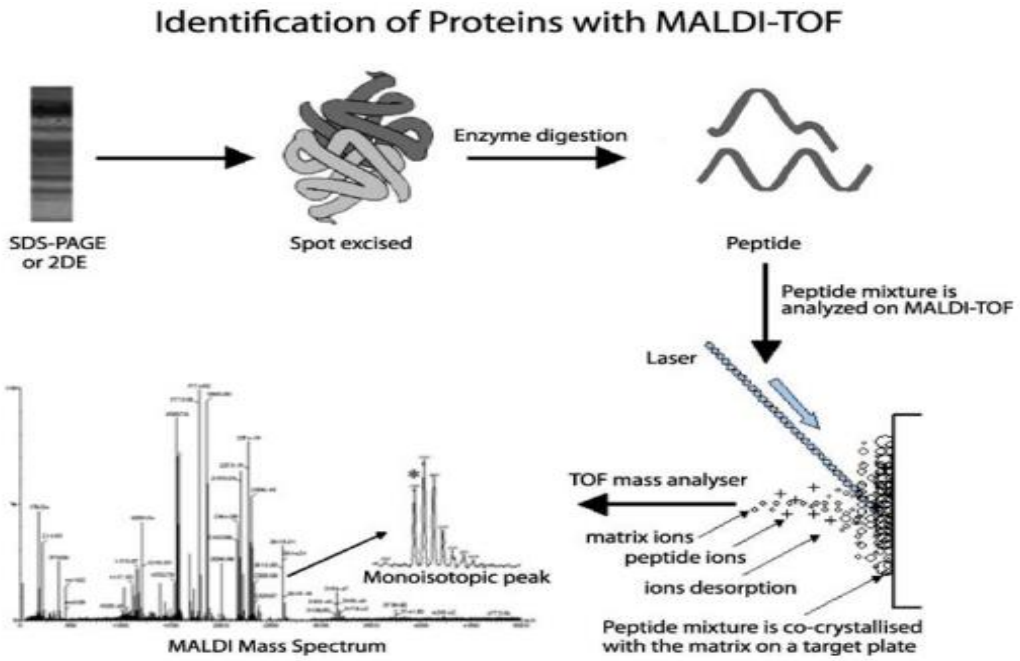

Figure 5. The MALDI ionization-MS nal transduction pathway for protein identification. In this study, the interplay of 32 identified and potential NF-alpha/NF-kappa B pathway components is examined using the TAP tag method. The various tagged proteins were stably transfected into HEK293 cells that respond to TNF-alpha. Both the non-induced and TNF-alpha-induced cells had their protein complexes isolated. One benefit of this study was that it was able to detect protein interactions in a variety of environments.

A Unique Perspective on FLAG Tag Immunopurification with a Single Tag

Ho et al. (2002) conducted a study that was similar to the large-scale examination of yeast protein interaction that was carried out in 2002 by Gavin et al. Examining 725 proteins, this study found 3,0617 interactions between 1,578 distinct proteins, which accounts for almost 25% of the yeast genome. Using anti-FLAG antibodies, IPs were carried out. To summarise, anti-FLAG-antibodies coupled to sepharose beads were used for immunopurification of cellular lysates obtained from cells transfected with the target FLAG-tagged protein. After that, the protein complexes were separated using 1-D SDS-PAGE and MS 73 stained with colloidal Coomassie dye to identify protein-protein interactions. After bands were removed from polyacrylamide gels, they were reduced and S-alkylated. Afterwards, they were hydrolyzed by trypsin. The next step was to do electrospray ionisation (ESI) coupled LC-MS/MS analysis on the digested peptides. To produce protonated molecules in the gas phase, another soft ionisation technique is employed, and it is called ESI [11]. Here, a volatile solvent is used to dissolve the analyte at a low concentration. Using a hypodermic needle and a high voltage, the analyte-containing solvent is pushed at a low flow rate to electrostatically distribute, or electrospray, tiny droplets that are micrometres in size. The analyte molecules absorb the charge from these droplets as they evaporate quickly. With electrospray ionisation, the sample's structure is maintained because the process takes place under air pressure. A nebuliser gas can be used to stabilise the spray. Then, for analysis, molecules are efficiently transported into MS. Using affinity purification in conjunction with electrospray ionisation to identify proteins is summarised in. One less complicated and more reliable method for finding protein interactors is FLAG immunopurification. In comparison to the original TAP, which is around 20 kDa, this tiny hydrophilic peptide has the added benefit of being about 1 kDa in size. Essential yeast proteins tagged at the C-terminus by 18% of the cases resulted in non-viable strains. There needs to be a reduction in the size of TAP tags due to the high percentage of nonviable strains, which highlights the significant advantage FLAG has over them [12]. However, when contrasted with the TAP tag, FLAG-IP has a greater rate of false-positive protein-interaction detection. The necessity for cross-validation of reported interactions arises from the significant false-positive rate in studies involving protein-protein interactions. Ewing et al. conducted a novel large-scale analysis of 338 protein-protein interactions in humans (2007). Researchers in this study examined protein-protein interactions in HEK293 cells using the FLAG-IP method in conjunction with ESILC-MS/MS [13]. By analysing these protein interactions, 6,463 interactions involving 2,235 distinct proteins were generated, with a further validation of 24,540 possible connections. A high level of confidence in the data quality was achieved when the dataset obtained using this method was validated using several methods. Arifuzzaman et al. (2006) used 4,339 His-tagged Escherichia coli ORFs in a large-scale pull-down study, one of several attempts to examine protein-protein interactions in lower species. The His-tag is purified on a nickel column, in contrast to FLAG-IP, which uses anti-FLAG antibodies coupled to sepharose beads. Then, MALDITOF MS was used to identify the purified proteins [14].

Purification of Phosphopeptides

Because proteomes are constantly changing, studying them is no easy task. Protein function can be modulated in part by post-translational changes, or PTMs. The addition of various PTMs has the potential to alter a protein's interaction partners by modifying its conformational structure. Among these critical PTMs, phosphorylation regulates numerous cellular activities. In order to determine which proteins interact with synthetic peptides of each cytosolic ErbB-receptor family member due to phosphorylation, Shulze et al. (2005) conducted a study. The 89 tyrosine residues that make up the cytosolic ErbBreceptor family were examined in this work together with 94 pairs of singly phospho-, non-phospho-, and doubly phosphopeptides. Immobilised on streptavidin-coated magnetic beads were peptides

containing tyrosine residues that had been synthesised with a desthiobiotin tag [15]. The protein complexes were analysed by MS after being eluted by biotin following treatment with cell lysates. Among the interactors identified were numerous well-known members of the ErbB-receptor family. The data also demonstrated that ERBB4 and the EGF receptor were more important signalling components than ERBB2 and ERBB3. The results of this investigation demonstrated the potential for large-scale peptide-protein interaction screens to provide light on the worldwide nature of interactions among protein families [16]

Mast Spectrometry Without Gel

The use of two-dimensional electrophoresis (2DE) in proteomic analyses has a long history. The complicated protein mixture was resolved and visualised using multiple forms of staining on 2DE gels [17]. The introduction of MS instruments capable of recognising the distinct spots in 2DE gels was a major step towards solving the main difficulty of protein identification. The demand for high-throughput approaches and the limitations of 2DE led to the development of gel-free alternatives. This section will centre on a method for determining protein-protein interactions known as multidimensional protein identification technology (MudPIT). It is common practice to use a strong cation exchange (SCX) column as the first chromatographic separation dimension in MudPIT, a gel-free peptide separation method that employs two chromatographic separation processes before sample ionisation and MS identification. The next step is to use reverse chromatography (RP) to separate the sample in a second dimension. RP has the benefit of efficiently desalting the sample mixture and being compatible with electrospray ionisation. Denaturation, reduction, and alkylation are the steps used to prepare samples. The sample is subsequently digested using the suitable enzyme. One last step before loading the samples onto the SCX column is acidification [18]. One benefit of MudPIT is that it can separate peptides in two dimensions: one for their charge state, and another for their hydrophobicity. Research by Graumann et al. (2004) demonstrates how MudPIT can be used to detect protein-protein interactions. The scientists analysed the interplay between twenty-one proteins that have a role in transcription and mitotic progression. After being TAP-tagged, the proteins of interest were purified and examined by MudPIT-MS. The authors found 279 possible physical interactions and 102 known ones using this method. Analysing materials using gel-free methods necessitates an in-solution digestion. Proteins in solution can be digested using a variety of techniques, including peptide fractionation using immobilised trypsin in monolythic columns or protein separation coupled to immobilised trypsin. Our group has recently released a novel in-solution digestion process. The proteomic reactor is a solitary microfluidic device that cleans, derivatizes, and digests proteins as part of this technique. The proteome reactor is a protein digesting system that uses a low-pH loading of cell lysates into a SCX column. Most peptides will have a positive charge under these conditions, which will help them bond to the reactor material. Trypsin does not work because of the acidic pH [19]. Raising the pH triggers trypsin, which in turn digests proteins. The next step is to use buffers that are suitable with MS analysis to elute the peptides. New gel-free proteomic approaches are being developed, which provides researchers with more instruments to solve a variety of biological problems.

Data on Protein Interactions and Their Quality

A wealth of information regarding the roles of several proteins that have not been identified can be gleaned via the creation of massive protein interaction datasets. However, validating these data sets is an incredibly challenging undertaking. Researchers now have the opportunity to compare data produced by diverse methodologies due to the increasing availability of large-scale data sets. Take, for instance, the comparison of data from the two additional Y2H experiments [20] with those from Gavin et al. (2002) and Ho et al. (2002). exhibits an extremely high rate of false positives, reaching up to 80%. The methods used are only one of several potential causes of the high false-positive rate. For instance, we have already established that Y2H has serious problems that raise its false rate. Incorporating a tag into the target protein changes its molecular blueprint. Adding a larger tag, such as the TAP tag, makes this effect more noticeable. To lessen this problem and achieve the least amount of interference with the protein structure, smaller tags can be alternately attached to the N- and C-termini of the protein. The process of coimmunopurification is one way that protein-protein interactions can be confirmed [21]. Due to the vast number of interactions in these research, validation is extremely difficult, if not impossible. Alternatively, large-scale data sets can be enhanced with the use of bioinformatic techniques. Ewing et al. (2007) conducted a large-scale experiment and used multiple strategies to guarantee high-quality results. Six distinct criteria were used to grade the data produced by each IP. Each protein-protein interaction was assigned a confidence score using this procedure, which allowed the authors to assess

its validity and establish a threshold for accepting the generated data. From a total of 6463 protein-protein interactions, 2251 received very high confidence scores after applying these criteria. Hart et al.'s (2007) unsupervised probabilistic scoring scheme is another illustration of this. In this method, we employ the matrix method of interpretation to create interaction data sets that contain all the prey-prey interactions from a specific bait pulldown. For each interaction, we assign a confidence score. This method can be used to combine data sets from different sources and boosts recall and/or precision compared to other ways, such as the traditional spoke model interpretation, which only takes bait-prey interactions into account. The data from Gavin et al. (2002), Krogan et al. (2006), and Ho et al. [22] were combined using this scoring scheme. that year (2002). According to the findings, the scoring metric outperforms the other groups' filtering systems. The development of consistent experimental protocols is a key area for future advancements in our understanding of protein-protein interactions. Specifically, this has to detail where the samples came from and how they were evaluated. To that end, there is an effort underway to establish proteomics workflow rules that will allow for easier data comparison, interchange, and verification; this effort is known as the Proteomics Standards Initiative (PSI). The "minimum information about a proteomics experiment" (MIAPE) recommendations are the aggregate name for these documents. With "minimum information about a molecular interaction experiment" in its name, MIMIx is one of the MIAPE modules. Following these rules should help users evaluate the reliability of the data, identify potential connections between their target proteins, and locate the primary sources that detail the experimental design. While these standards may not yet cover all bases when it comes to standardising protein interaction data, they do lay the groundwork for future, more comprehensive guidelines.

Enzyme-Protein Interaction Registry

A public database of all interactions between proteins is necessary due to the generation of enormous interaction data sets. The results of many different kinds of protein-protein interaction investigations are now available in a number of databases. Some examples of such databases are BioGRID, BIND, DIP, IntAct, MINT, and MPact (MIPS). as well as HPRD. Data sharing was a challenge in the beginning since these databases were operating independently, lacking uniform extraction, curation, and storage processes, and because they did not all investigate the same scientific articles. The data was often more complementary and could be combined to enhance our understanding of interactome networks, as we found that only a tiny portion of those databases were overlapping. A variety of protein interaction databases have recently banded together to form a special effort that demands a community-standard for how data about protein interactions is represented. The Molecular Interaction (MI) subgroup of the PSI is responsible for developing this model. Already, this type of work has been embraced by major interaction databases. The International Molecular Interaction Exchange (IMEx) consortium was founded by a number of databases that have used these guidelines as a roadmap. IMEx is now working on a system to share curation tasks and eventually exchange curated data .

Human disease research using mass spectrometry-based protein-protein interaction networks

In order to effectively prevent, diagnose, and cure human diseases, it is essential to identify their principal molecular foundation. Scientists have put a lot of faith on massive genome research to help them understand disease pathways in the last 20 years. However, the molecular mechanism of the majority of diseases is still a mystery, even though a great deal of genetic data has been collected. Many human diseases are complicated and do not adhere to a traditional genotype-to-phenotype model, which helps to explain this phenomenon. A pathogen infection, epigenetic alterations, or a series of mutations can all lead to these conditions. In cancer, when a unique set of mutations is not always unique to a particular form of cancer, the error of assuming basic genetic changes to explain complicated disease phenotypes is most clearly shown. Furthermore, diseases can result from mutations in a single gene since the proteins encoded by that gene have many purposes in various biological settings (Nadeau, 2001). As a result, it might be challenging to derive valuable diagnostic or prognostic information only from genetics. To get around this problem, we need to think about genetic data in relation to broken down cellular processes and networks. Understanding the impact of disease mutations on a complex network of interrelated pathways is best accomplished using systems biology approaches. These approaches strive to present a complete picture of a biological process by quantifying all observable components and their interactions. These networks rely on proteins as their central component. Proteins often rely on interactions with other proteins to carry out their functions rather than operating alone. Therefore, a potent approach for determining the functional effects of genetic diversity is the analysis of protein-protein interaction (PPI) networks.

Important protein-protein interactions (PPIs) in cellular processes are associated with disease-related gene alterations in this method [23]. By comparing disease states with the wild-type reference map, we can learn more about how networks change during disease pathogenesis. This can be achieved through the introduction of mutated proteins or the exogenous production of pathogen proteins. The process of adapting to alterations brought about by diseases is mostly carried out by cellular proteins. A disease-related mutation can affect more than just one gene product due to the interconnected nature of proteins. Rather, it influences the full network, which in turn can influence the activity of a subset of proteins. Analysis based on PPIs look at the pieces of pathways rather than the genes or loci that are thought to be responsible for human disease. that are most affected by the illness, providing an alternate way to determine how a mutation affects cellular function. In a PPI study, the "bait" proteins are represented by nodes in a network-based method to protein interaction visualisation. Protein interactions detected by Affinity Purification Mass Spectrometry (AP-MS), proximity labelling, Crosslinking Mass Spectrometry (XL-MS), or other investigations are linked to nodes through edges. In both the diseased and non-diseased (WT) phases, this mapping is carried out, and alterations in the global control of PPIs in the networks are tracked. Disruptions to these networks, such as the elimination of connections altogether, the reduction of certain interactions to a smaller subset, or the addition of novel interactions, might result from the introduction of disease-related mutations. Due to the interconnected nature of PPI networks, even very little perturbations, like the insertion of a single gene mutation, can have far-reaching effects at various nodes in the system. An infection or changes in the disease-related protein's interaction partners might cause a particular disease state by bridging the gap between genotype and phenotype. There are several therapeutic and clinical benefits to studying human diseases using a network-based approach. If a gene or protein is found to be involved in a specific biochemical process or disease, it's likely that its interacting proteins are also involved. This means that there may be more than one protein or gene involved, and that there may be potential therapeutic implications and mechanistic explanations for these processes. In this review, we take a look back at how far we've come in our quest to understand human disease utilising impartial PPI networks based on mass spectrometry (MS). We will focus on the present state of the field and how recent developments in PPI network mapping have helped to alleviate some of the problems that have been identified. To learn more about alternatives to MS-based PPI identification methods, read on [24].

MS-based approaches for worldwide PPI research

Protein quantification using liquid chromatography-MS (LC-MS) is a selective, accurate, and sensitive technique. The universal and unbiased character of MS proteomics is one of its primary benefits when it comes to identifying PPIs. The yeast-2-hybrid (Y2H) approach, for example, maps the physical, binary interactions of a set of target proteins. This is different from other approaches to PPI identification.

Scientific investigations using affinity purification mass spectrometry (AP-MS). employ epitope tagging, which entails fusing a small peptide or protein tag (such as FLAG-, TAP-, Strep-Tag, or cmyc) to the target protein. This fusion can take place within an external expression construct or at the gene's endogenous promoter using gene editing tools like CRISPR-Cas. In contrast to lower throughput immunoprecipitation (IP) investigations, the generated bait protein acts as an affinity capture probe for interacting, or "prey" proteins. This eliminates the need for particular antibodies to the target proteins. An epitope-recognizing matrix makes affinity tag purification a breeze. In order to identify interacting proteins, MS must be performed after washing processes to remove non-specific interactors. Thanks to developments in high-throughput AP-MS methods, extensive interaction networks in both healthy and disease models have been able to identify thousands of protein complexes and PPIs. To far, the BioPlex database has collected more than 56,533 interactions with 10,961 proteins in HEK293T cells, making it the largest assemblage of such PPI networks. Important tools for biomedical research, like as publicly available data sets like hu.MAP 2.0, have led to several discoveries of disease-underlying molecular pathways, some of which are discussed here. The requirement for less harsh lysis conditions compared to those normally used in MS investigations is a drawback of AP-MS. Protein extraction issues make this method unsuitable for capturing membrane proteins. It is also possible to lose weaker or more fleeting connections during washing or extracting. To improve the recovery rate of proteins lost in standard AP-MS research, tandem affinity purification (TAP) tagging attaches two distinct proteins or peptide tags to an interest fusion protein. One tag, such as His-tag, can withstand harsher lysis or washing conditions. The downsides of this approach include the need for more time-consuming sample preparation and purification and the

possibility of artefacts caused by the insertion of big tags to the target protein. Negative controls must be carefully chosen since background difficulties can occur regardless of the amount of tags used because non-specific interactors that persist after washing still exist. Another drawback of APMS is that it can lead to false positive PPI identifications due to the mixing of normally inactive cellular compartments caused by lysis. In the section on New Methodology, we will talk about some of the techniques that are being considered right now to deconvolute the consequences of compartment mixing [25]. It is beneficial to test tagging both ends of the protein because adding a tag to either the N- or C-terminus could interfere with its natural activity. Keep in mind that AP-MS isn't great at telling direct interactors apart from indirect ones. However, AP-MS has numerous benefits over previous methods for identifying interactions (such as Y2H), such as high sensitivity and the simultaneous quantification of non-binary interactors. Interaction protein post-translational modifications (PTMs) can also be detected using AP-MS. An intensity value for a specific protein can be obtained by label-free measurement after data acquisition. With these numerical values, we can do comparison analysis, which should help us find out if an interaction is protein specific or not.

Distance classification

In addition to the standard AP-MS investigations, proximity labelling can be used. In this scenario, a promiscuous labelling enzyme is linked to a bait protein of interest and expressed in cells to monitor proximal proteins. Endogenous proteins can be covalently tagged within a 10-20 nm range by adding a small molecule substrate, such biotin. This captures the protein's surroundings, including potential interactors.

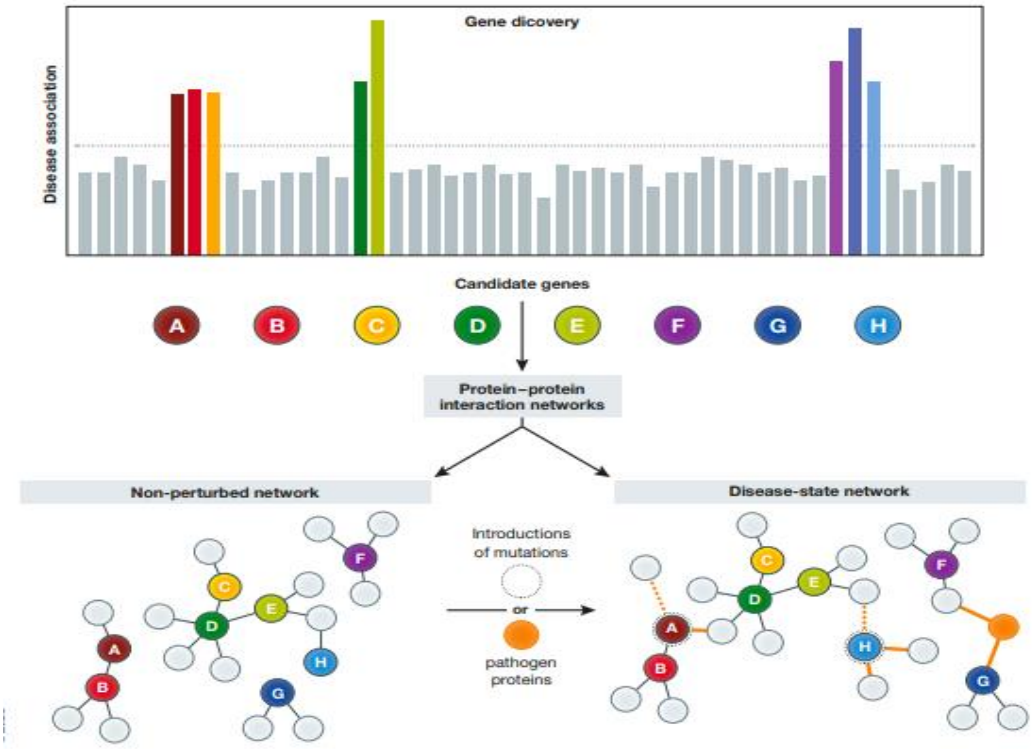

Figure 6. A systemic method for translating genomic data into a pathway-level knowledge of the data. The foundation of PPI networks can be utilised by genetic variations, which may be uncommon among individuals with a particular condition. To better understand the functional impact of disease-related mutations, it can be helpful to compare WT and PPI networks that have had these mutations added. A similar mechanism at work during infection is the introduction of pathogenic proteins, which can dictate which host pathways are hijacked.

Proximity labelling allows for more efficient protein extraction, lysis methods, and harsher washing conditions than AP-MS, allowing the identification of weak or transient interactions that might be lost with other methodologies. This process begins with cell lysis, where proteins are denatured and solubilized. Then, biotinylated proteins are selectively enriched, usually through streptavidin binding. Finally, MS is used for identification. Detergents are used during lysis in the technique because complexes do not need to be intact during lysis and purification. We have devised a number

of proximity labelling technologies. BirA, a biotin ligase that has undergone certain mutations that make it promiscuous, is utilised by BioID. Proteins in close proximity to BirA undergo covalent biotinylation when the biotin cloud that forms undergoes a biotin-to-reactive-form change. Nuclear envelope, centrosome, nucleus, cytoplasm, Golgi apparatus, endosome, lysosome, mitochondrial matrix, cell-cell junctions, and flagella are all subcellular compartments that BioID has targeted; however, labelling efficacy is limited in the ER. Labelling must be carried out for 18–24 hours in BioID for enough material to be identified by MS, because reaction kinetics are slow. The standard procedure for discovery MS involves breaking a protein mixture into peptides with specific cleavage sites (e.g., with trypsin), which are then separated by liquid chromatography and their mass-to-charge (m/z) is measured in a mass spectrometer. By gathering a second MS spectra following induced fragmentation, the sequence of individual peptides can be ascertained in conventional tandem MS/MS investigations. We computationally search vast databases specific to the organism of interest using the combined m/z data of fragments and complete peptides. This process identifies proteins in the original mixture. Data will be "scored" to ascertain the veracity of the identified interaction in order to identify potential interactors in investigations of protein-protein interactions. A number of factors, including the quantity, specificity, and repeatability of each discovered protein, are typically combined to achieve this. For this reason, there are a number of scoring algorithms available, such as SAINT, CompPASS, and MiST. There are some key differences in the approach taken by the two algorithms. SAINT uses prey-specific quality controls and quantitative data to calculate the likelihood of a true positive interaction between bait and prey proteins, while COMPPASS uses a number of scoring parameters that centre on abundance, uniqueness, and reproducibility to differentiate between genuine interactors and background contaminants. There are a number of MS methods available for quantifying changes between circumstances, and there are also computational ways that compare PPIs to suitable controls to determine their specificity. Countless samples can be used to compare the relative abundances of identified proteins using label-free quantification. One drawback of this method is that it requires injecting the same amount of each sample onto the analytical column in order to make meaningful comparisons. Data normalisation becomes necessary when this is not an option. It is recommended to analyse the samples being compared in a single acquisition batch on the mass spectrometer in order to minimise instrumental bias. Avoiding systemic errors can also be achieved by randomising the run order. Users can increase experimental throughput by multiplexing numerous samples together using metabolic or isobaric labelling techniques like SILAC and tandem mass tag (TMT), among others. Stable heavy amino acids are biologically incorporated into proteins by SILAC. Chemical tags are applied to free amines in vitro after digestion via isobaric tagging procedures, which make use of molecules triggered by the NHS. In order to determine the source of a certain protein interactor in a mixture of control and experimental samples, all labelling approaches depend on adding a mass label to extra control samples. All things considered, these techniques make it possible to compare results from various settings or times, or to separate out interactions that are specific from those that are not. For more consistent, sensitive, and accurate interaction validation, focused MS procedures like multiple/selective reaction monitoring (MRM/SRM) or parallel reaction monitoring (PRM) are available. In a nutshell, when developing an assay, certain peptides from the target protein are chosen. The last experiment involves monitoring these using their unique fragment ions to ensure accurate quantification. Many non-specific interactors or contaminants are copurified with the protein of interest among the identified proteins in MS-based interaction studies. That is why it's critical to distinguish between real interactors and artefacts when analysing PPI investigations. A combination of well-planned experiments and appropriate controls can help achieve this goal. To establish the specificity of interaction, suitable controls must be utilised, such as a similar protein with the same tag or the tag alone. In control trials, for instance, GFP can serve as a bait. Because of the nature of the epitope tags or the affinity capture method, it is quite probable that the detected interactors are false positives, and it is highly improbable that GFP forms interactions with many proteins. Furthermore, distinct affinity tags can detect different types of background contaminants. The CRAPome database, which is a public collection of interactions derived from negative control data, can be used to retrieve and filter out these contaminants from trials. Another kind of contamination is the carryover of overexpressed proteins. This happens when a small amount of protein is detected in later MS tests, even if it isn't actually an interactor. It is possible that this issue can be resolved by implementing rigors wash procedures between different testing situations. limits the kinds of studies that can be done with BioID. Because of the constraints of time, BioID trials can only provide static interaction maps. The Aquifex aeolicus biotin ligase was engineered with mutations to produce BioID2, a bioID alternative. This much smaller enzyme improves targeting and

localisation to subcellular compartments by reducing disruption to the fusion protein. But labelling still takes more than sixteen hours. Two enzyme variants, TurboID with fifteen mutations and miniTurbo with thirteen mutations and a loss of the N-terminal domain, were produced by directed evolution of BirA, which Branon et al. (2018) used to increase the efficiency and speed of labelling. Labelling that is on par with BioID can be accomplished in less than 10 minutes thanks to the strong biotin affinity of these enzymes. Peroxidases are enzymes that catalyse redox processes; a new class of proximity labels emerged from these alterations. One of the peroxidases used for proximity labelling is horseradish peroxidase (HRP), which has been the subject of the most research. But it's not very efficient at reducing environments when it comes to labelling (Trinkle-Mulcahy, 2019). This limitation is not present in engineered ascorbic acid peroxidase (APEX), which can be genetically inserted as a tag onto target bait proteins (Rhee et al., 2013; Hung et al., 2016). Biotin labelling kinetics on the order of minutes are provided by the covalent reaction between biotin-phenoxyl radicals, which are formed when phenol derivatives are oxidised by APEX after the controlled addition of H2O2. This method is ideal for studying fleeting or ever-changing protein interactions because of APEX's fast labelling capabilities, which are on par with many biological processes. Since APEX remains active in reducing settings, such as the cytosol, labelling can be carried out in a wide variety of subcellular contexts. However, since peroxide inhibits APEX labelling in live creatures and may have detrimental effects on cells, its necessity has been questioned. In order to minimise labelling times and possible toxicity concerns, more recent versions of proximity labelling methodology have been developed. A new contact-specific technology called SplitID splits the TurboID enzyme into inactive pieces. Like interacting proteins, these two pieces recombine when they're near each other. When using this procedure, off-target labelling can be eliminated because biotinylation is limited to the contact sites between fragments that have been targeted to certain organelles. Just to how peroxidase activity is enhanced when the N- and C-terminal portions of split APEX are connected through molecular contacts, they are inactive when separated. Prior to conducting a proximity labelling experiment, much thought should be given to the experimental design. Regardless of the proximity labelling method used, all of them catch proteins that are close to the bait. It might be challenging to differentiate proteins that truly exist in the near vicinity when proteins that are not direct interactors yet colocalize during the labelling phase do so merely as a result of diffusion across the enzymatic labelling zone. This phenomenon is known as high background. Unexpected interactor proteins can be located by doing parallel analyses of expressed ligases that do not have bait proteins attached. Proteins in this control sample may have attached to the streptavidin utilised for enrichment or formed naturally by interacting with the ligase. Enzyme insertion at the N- or C-terminus of a protein may change its function, just as it does in AP-MS. Make that normal localisation is not disrupted by testing enzymatic fusion on the N- and C-termini of the protein of interest before creating an enzyme-expressing stable cell line. It is also possible that proteins close to the unlabeled end will go undetected since they are outside the labelling radius. Therefore, it could be beneficial to do independent tests that label the N-terminus and C-terminus .

MS with crosslinking capabilities

While AP-MS can determine which proteins are part of a complex, it cannot tell you which ones are physically touching each other. This void can be filled by XL-MS. It reveals the structure by locating pairs of nearby amino acids that are covalently bound by a chemical crosslinker of a given length; this includes weak or transitory contacts. The complex and PPI binding interface architecture and subunit orientation can be inferred from the distance restriction data collected. Almost ambient conditions are used to carry out the crosslinking process. First, enzymatic digestion is used to generate cross-linked peptides. Then, these peptides are enriched. After that, MS analysis and database searching are used for identification. The examination of the data reveals details about the localisation of particular cross-linked amino acid residues and the sequence assignment of cross-linked peptides. The physical interaction data obtained by XL-MS can be utilised to enhance computational modelling and structural biology research when combined with integrative modelling approaches. Several protein complexes and interactions across the proteome have their structures deduced from the restraint information provided for both intra- and inter-linked proteins (Shi et al., 2014). The data analysis step is when XL-MS falls short. There are a lot of different kinds of cross-linked peptides in spectra, and it's not easy to account for every potential combination of them all. However, programmes that can aid in this process have been developed at a rapid pace, with versions tailored to cleavable and non-cleavable cross-linkers. There is a wide selection of cross-linkers that can be used to target specific amino acid side chains and binding interface distances. Numerous approaches for enriching cross-linked peptides, taking advantage of the benefits of diverse cross-linker chemistry, and optimising data gathering and processing workflows have been developed to

enhance the sensitivity and accuracy of XL-MS. To stabilise in vivo contacts and make them resistant to denaturing and washing, which remove background contaminants but preserve weak or temporary interactions, crosslinking must be done before cell lysis. For interactome mapping, the aforementioned methodologies offer complementary approaches .

The establishment of health information systems

In the beginning, researchers employed AP-MS to find the proteins that would bind to 338 bait proteins that were chosen for their involvement in human disorders like cancer, diabetes, and obesity. After a thorough examination of the data, the authors found 6,463 interactions involving 2,235 proteins and proved that bait proteins had a propensity to attract partners with similar functions. Several small- and large-scale investigations have utilised AP-MS and proximity labelling since then to uncover the interaction partners of proteins implicated in various illnesses. The disease mechanisms of certain mutations can be identified by directly comparing interaction networks and how they alter in response to perturbations associated to disease. One study that employed this method to confirm the causal involvement of a hereditary mutation in the RAS oncogene family-like 3 (RABL3) and its strong association with familial pancreatic cancer was Nissim et al. (2019). Until recently, RABL3's biological role remained a mystery. By comparing the interactomes of wild-type and mutant RABL3, researchers were able to identify RABL3's roles as a regulator of KRAS and a potential biomarker for pancreatic cancer. Another study that demonstrated the usefulness of comparing disease-related PPI networks with healthy ones was the one conducted by the Yates group. They employed IP-MS to identify the interactions that drive the CF phenotype. Deletion of phenylalanine 508 from the CFTR gene, which leads to protein instability, is the main cause of cystic fibrosis (CF), a Mendelian condition. A changed PPI network, characterised by the insertion of new interaction partners not found in the WT CFTR network, is created when the ΔF508 CFTR misfolds and recruits different chaperones. Between wild-type and mutant CFTR, a total of 209 interacting proteins showed substantial differences in abundance, number of nodes, or both. We analysed the changes in the interaction partners between WT and ΔF508 CFTR under conditions known to promote CFTR glycosylation, such as a shift to lower temperature and inhibition of histone deacetylase, because glycosylation of ΔF508 CFTR can partially restore its function. There were notable differences in the WT and ΔF508 CFTR interactomes under both circumstances. As an illustration, when the temperature is set to 30°C during incubation, 89% of the unique interactions with ΔF508 CFTR are eliminated. These interactions include those with heat-shock proteins that are involved in protein folding, RNA-processing proteins, rescued proteins that are involved in ER quality control, and degradation proteins of ubiquitin-mediated pathways and ERAD. Neurodegenerative diseases, other disorders caused by protein misfolding or aggregation, and numerous proteins exhibiting differential interactions between WT and ΔF508 CFTR were linked, which raises the possibility of shared pathways and mechanisms among these seemingly unrelated conditions (Pankow et al, 2015)

Blending orthogonal datasets

The finding potential can be further enhanced by integrating MS-based PPI interactomes with complementing data. The functional interactions between genes can be better understood, for instance, with the use of genetic interaction data. Researchers have been able to compare thousands of genes associated with different biological processes through genetic interaction studies, which map the phenotypic output between pairs of reduced genes. Novel treatment approaches targeting synthetic lethal partners of inactivated tumour suppressors can be developed using this technique, which reveals which genes are cooperating. Crucially, data regarding functional links can show molecular cross-talk between pathways in addition to identifying direct physical interactions. For instance, by combining GI research with PPI networks, new b-catenin regulators were found, and the functional networks needed for b-catenin-active tumours to survive were defined. To study the basic genetic relationships of HIV-1 infection, this method was only lately modified. Over 63,000 comparisons were made as a result of the pairwise depletion of 356 HIV-1-related human genes, including host-dependency characteristics found in an earlier AP-MS investigation. This study found that the RNA deadenylase complex CNOT, which had never been linked to this process before, is involved in the innate immune response's mediation of HIV-1 infection. One more potent option is to integrate data across networks of various diseases. Investigating the function of viruses in carcinogenesis is one area that has made use of network analysis. Since viruses can cause a variety of malignancies, it's reasonable to assume that tumours and viruses share some genetic ground, even though the specific genes they modify are distinct. Research has utilised AP-MS to identify

PPI networks spanning tumor-associated proteins from several viral species, with the goal of elucidating the function of viral proteins in this process. All HPV viral proteins were subjected to a more targeted examination. Finding cancer genes in these networks brought attention to the fact that several viruses associated with cancer rewire Notch signalling. However, by incorporating network propagation to analyse differential mutation patterns in HPV-associated cancer samples, we were able to identify oncogenic activities mimicked by the virus, such as enhanced invasion of tumour cells that depended on the interaction of viral and human proteins. When combined with structural biology methods, MS-based interaction data can shed light on the structural alterations that happen as diseases evolve. X-ray crystallography, cryo-EM, and XL-MS can be used to map disease-related protein complexes, which can then be used to guide the discovery of therapeutic drugs. A well-researched virus that causes lethal epidemics repeatedly, the Ebola virus (EBOV) infection was examined using this method. Among the seven EBOV proteins, six had virus-host interactomes that uncovered 194 high-confidence interactions. Among them, one included the viral VP30 protein and the human ubiquitin ligase RBBP6. By simulating the binding of the EBOV nucleoprotein (NP) to VP30 at the same interface, X-ray crystallography has shown that this contact blocks the virus-virus PPI necessary for viral transcription. Thus, RBBP6 is a host restriction factor, and peptide mimics effectively blocked the replication of EBOV in cell cultures. Based on these results, it appears that inhibitors could be useful in targeting the VP30-RBBP6 binding interface. In order to determine the orientation of subunits and their binding sites that are involved in disease aetiology, other groups have employed comparable methodologies. The crystal structure of the Chlamydia complex was determined using cryo-EM and AP-MS, which revealed interactions between the human proteome and the intracellular pathogen Chlamydia trachomatis. This study provides new information about retromer assembly. Despite the fact that cryo-EM allows for protein structures to be resolved down to the atomic level, there are some configurations that are prone to low electron density regions, which makes it challenging to resolve individual subunits in the absence of orthogonal structural data. The cryo-EM volume can include potential crystal structures, with the subunits' locations and orientations determined by the distance constraints imposed by XL-MS studies. Using this method, Henry et al. determined the binding processes and active structure of apolipoprotein E4 (ApoE4). ApoE4 has been associated with Alzheimer's disease and has a role in lipid transport. The resolution of two ApoE4 conformations allowed researchers to deduce that the accessibility of the receptor-binding region determines the activation mechanism of ApoE4. Collecting illness networks has the potential to enhance our comprehension of fundamental biological processes in general and the disease model being studied in particular .

Novel breakthroughs

Completing an interactome is difficult because protein-protein interactions are dynamic and can change depending on factors like stress, time, environmental stimuli, disease status, and cell line or tissue type. Our knowledge of disease networks and the development of new high-throughput technologies, MS methodologies and modifications, and novel combinations of current approaches are dependent on these developments. In order to manipulate a biological system for the purpose of illness treatment, it is necessary to have a thorough fundamental understanding of that system. This can be achieved through the adaptation of modern proteomics techniques. Progress is being made in the sector mainly due to quantitative MS methods, which include targeted approaches and improved acquisition techniques. The examination of AP-MS samples has revealed that data-independent analysis (DIA) is the most promising of these methods. An effective method for acquiring MS/MS data, DIA seeks to quantify all peptides expressed in a specific proteome. All peptides within a specific m/z range are fragmented, enabling data to be collected for all peptides in a mixture, instead of sampling the most intense precursor ions present in an MS1 scan, as done in typical data-dependent techniques. Iteratively fragmenting the full m/z range is achieved by this method. In comparison to conventional data-dependent acquisition (DDA), DIA can enhance sample reproducibility by consistently fragmenting the same subset of peptides between trials, which in turn reduces the number of missing results. Furthermore, DIA approaches allow unbiased analysis in addition to providing very accurate and reproducible quantification, comparable to targeted proteomics methods. Improved PPI network descriptions are on the horizon thanks to the enhanced dynamic range and sensitivity of these new MS approaches, which reduce false negatives when investigating protein-protein interactions. In addition, they make it possible to quantify the identification of interaction modifications. In order to score changed interactors following medication exposure or related to illness status, Lambert et al. utilised DIA to build a pipeline. The human kinase CDK4 was analysed using this pipeline to measure the variations in interactors between the wild-type and melanoma-associated sequence variants. By capturing known and unknown

interactors, DIA showed that the recruitment of HSP90 to CDK4 mutants at Arg24 is specific. Not only do conventional approaches have sensitivity difficulties, but sample sets can also grow to unmanageable sizes very quickly. All the point mutations in a single gene that cause a disease can be hard to identify and study on an individual basis. It could become tedious to extend the investigation to multiple genes or circumstances. In an effort to find a solution to this issue, researchers have looked into ways to detect PPIs worldwide without using affinity purification. For experiments that aim to characterise changes throughout the full interactome, these methods can be useful supplements to more conventional AP-MS procedures. To systematically identify protein complexes, Aebersold and colleagues have developed the idea of protein correlation profiling using complexcentric analysis. Here, size exclusion chromatography is used to separate protein complexes. Following proteolytic digestion, each fraction is examined using DIA MS and a recently created software programme called CCprofiler. The presence of proteins in the same fraction allows one to infer interactions. They discovered 462 protein complexes with 2,127 protein subunits in the HEK293 cell line during a proof-of-principle experiment. Without epitope tagging, this method shows tremendous promise for large-scale PPI identification. However, a high-throughput approach to tracking PPIs throughout the proteome was developed by Nordlund et al. using their cellular thermal shift test, which they called thermal proximity coaggregation (TPCA). A key premise of TPCA is that specific temperatures cause the denaturation of various proteins. One way to create a melting curve is to examine the rate of solubility loss due to denaturation at certain periods in time. An advantage of this technology is that interactions may be observed in vivo, as individual proteins denature, so any interacting proteins will have similar solubility profiles and can be analysed by MS. Preheating intact cells to the target temperature before lysis preserves preexisting contacts and eliminates post-lysis spurious interactions. Utilising sample multiplexing with thermal gradient tracking (TMT), this method can combine as many as sixteen samples subjected to varying temperatures. weather conditions. Melting curves for 7,693 proteins isolated from K562 cells and 111,776 PPIs that were made publicly available were compared. It was shown that pairs of proteins known to interact were statistically more likely to have similar melting curves than pairs of proteins drawn at random. Following this, melting curves were produced for a number of different cell lines in order to determine which compounds were common and which were specific to each line. Researchers employed this method to examine cancer medication-treated cells, finding not only the predicted binding partners but also numerous new interactors and downstream effectors. One typical argument against MS-based methods is that, when cells are lysed, their compartments are mixed up, and crucial data about the subcellular localisation of target proteins is lost. In light of this, the field of spatial proteomics has emerged as a catch-all for the adaptation of various MS methodologies in light of recent technical breakthroughs. These novel approaches allow for the quantitative evaluation of global protein translocation in both healthy and diseased states through high-throughput methods. One application of APEX proximity labelling is the extraction of subcellular localisation information from so-called "bystander proteins" that are present in the same labelling environment as the target protein but do not interact with it. Using this method to give spatial and temporal resolution during the fast signalling cascade of G protein-coupled receptor signalling was demonstrated in a proof-of-principle research. To better understand disease networks, both new and old MS methodologies are adapting. To better understand how adrenaline stimulates cardiac function, for instance, APEX was recently modified to mark functional, undamaged mouse hearts. The research of PPI networks and the development of new treatment approaches will be further advanced by the introduction of new technology and by combining current methods for protein purification and detection with new ones.

CONCLUSION

The ability to collect extremely large datasets and put disease-related mutations and changes in a biological context has been substantially enhanced by advancements in MS instrumentation, processes, and data analysis. A genetic mutation affects more than just one gene product due to the interconnected nature of proteins. As a matter of fact, it influences the function of entire protein subsets by operating on an entire network. By studying these interactions objectively, we can learn more about the temporal and spatial locations of molecules and the pathways that are impacted by various diseases. It would be excellent to place complexes into pathways and provide clues on their mechanisms by functionally analysing these data using genetic interaction maps or further integrating them with structural information from cryo-EM or XL-MS. Nevertheless, disease networks should not be considered independently. The importance of the mechanisms suggested by an interactome must be assessed, however, in the same way as with any systems biology approach. Because of its scalability, immortalised cell lines are a common tool

for PPI research. Though simple to work with, these cell lines may miss the mark when it comes to capturing relationships that matter in more complicated tissues and creatures. Modern genetic techniques, such as genome editing of primary cells using CRISPR/Cas9, can pave the way for the creation of more realistic models of human physiology. to investigate interactions using illness models that are relevant to function. We are getting closer to incorporating these technologies into personalised medical applications as technology keeps getting better and these methods become more widely available and have higher throughput. In addition to helping doctors understand how the body works, they may also pinpoint exactly where a patient's network is most vulnerable and advise them on the best courses of treatment. Proteomics has progressed greatly since MS was used for protein identification. This area will see further advancements with the introduction of novel affinity purification methods and MS machines. The rate of false positives and negatives can be reduced through more stringent experimental design and data processing. At long last, everything is in place to go on with the human interactome identification. On a grand scale, scientists will one day be able to examine how protein interactions change in response to various stimuli and compare the interactome of cells in various disease states or after treatment with various cues .

References

1. Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. Nature 537: 347 – 355
2. Antonicka H, Lin Z-Y, Janer A, Aaltonen MJ, Weraarpachai W, Gingras A-C, Shoubridge EA (2020) A high-density human mitochondrial proximity interaction network. Cell Metab 32: 479 – 497.
3. Banerjee S, Wei T, Wang J, Lee JJ, Gutierrez HL, Chapman O, Wiley SE, Mayfield JE, Tandon V, Juarez EF et al (2019) Inhibition of dual-specificity tyrosine phosphorylation-regulated kinase 2 perturbs 26S proteasomeaddicted neoplastic progression. Proc Natl Acad Sci USA 116: 24881 – 24891
4. Batra J, Hultquist JF, Liu D, Shtanko O, Von Dollen J, Satkamp L, Jang GM, Luthra P, Schwarz TM, Small GI et al (2018) Protein interaction mapping identifies RBBP6 as a negative regulator of Ebola virus replication. Cell 175: 1917 – 1930.
5. Beltran PMJ, Jean Beltran PM, Federspiel JD, Sheng X, Cristea IM (2017) Proteomics and integrative omic approaches for understanding host– pathogen interactions and infectious diseases. Mol Syst Biol 13: 922
6. Bouhaddou M, Eckhardt M, Chi Naing ZZ, Kim M, Ideker T, Krogan NJ (2019) Mapping the protein-protein and genetic interactions of cancer to guide precision medicine. Curr Opin Genet Dev 54: 110 – 117
7. Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY (2018) Efficient proximity labeling in living cells and organisms with TurboID. Nat Biotechnol 36: 880 – 887
8. Buljan M, Blattmann P, Aebersold R, Boutros M (2018) Systematic characterization of pan-cancer mutation clusters. Mol Syst Biol 14: e7974
9. Calderone A, Licata L, Cesareni G (2015) VirusMentha: a new resource for virus-host protein interactions. Nucleic Acids Res. 43: D588 – D592
10. Chang I-F (2006) Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes. Proteomics 6: 6158 – 6166
11. Cheerathodi MR, Meckes Jr DG (2020) BioID combined with mass spectrometry to study herpesvirus protein-protein interaction networks. Methods Mol Biol 2060: 327 – 341
12. Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America 95:14863–14868. doi: 10.1073/pnas.95.25.14863
13. Grallert A, Chan KY, Alonso-Nun˜ ez ML, Madrid M, Biswas A, Alvarez-Tabare´ s I, Connolly Y, Tanaka K, Robertson A, Ortiz JM, Smith DL, Hagan IM. 2013. Removal of centrosomal PP1 by NIMA kinase unlocks the MPF feedback loop to promote mitotic commitment in S. pombe. Current Biology : CB 23:213–222. doi: 10.1016/j. cub.2012.12.039

14. Haruki H, Nishikawa J, Laemmli UK. 2008. The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Molecular Cell 31:925–932. doi: 10.1016/j.molcel.2008.07.020

15. Hepp MI, Alarcon V, Dutta A, Workman JL, Gutie´ rrez JL. 2014. Nucleosome remodeling by the SWI/SNF complex is enhanced by yeast high mobility group box (HMGB) proteins. Biochimica Et Biophysica Acta 1839: 764–772. doi: 10.1016/j.bbagrm.2014.06.014

16. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. Nature 425:686–691. doi: 10.1038/nature02026

17. Hung MC, Link W. 2011. Protein localization in disease and therapy. Journal of Cell Science 124:3381–3392. doi: 10.1242/jcs.089110 J

18. oglekar AP, Bouck DC, Molk JN, Bloom KS, Salmon ED. 2006. Molecular architecture of a kinetochoremicrotubule attachment site. Nature Cell Biology 8:581–585. doi: 10.1038/ncb1414

19. Lambert JP, Mitchell L, Rudner A, Baetz K, Figeys D. 2009. A novel proteomics approach for the discovery of chromatin-associated protein networks. Molecular & Cellular Proteomics : MCP 8:870–882. doi: 10.1074/mcp. M800447-MCP200

20. Ledesma-Ferna´ndez E, Thorpe PH. 2015. Fluorescent foci quantitation for high-throughput analysis. Journal of Biological Methods 2:22. doi: 10.14440/jbm.2015.62

21. Lubbock AL, Katz E, Harrison DJ, Overton IM. 2013. TMA Navigator: Network inference, patient stratification and survival analysis with tissue microarray data. Nucleic Acids Research 41:W562–W568. doi: 10.1093/nar/ gkt529

22. Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM, Krogan NJ. 2008. Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science 322:597–602. doi: 10.1126/science.1162790

23. O´ lafsson G, Thorpe PH. 2015. Synthetic physical interactions map kinetochore regulators and regions sensitive to constitutive Cdc14 localization. Proceedings of the National Academy of Sciences of the United States of America 112:10413–10418. doi: 10.1073/pnas.1506101112

24. Ramey VH, Wong A, Fang J, Howes S, Barnes G, Nogales E. 2011. Subunit organization in the Dam1 kinetochore complex and its ring around microtubules. Molecular Biology of the Cell 22:4335–4342. doi: 10. 1091/mbc.E11-07-0659

25. Reid RJ, Gonza´ lez-Barrera S, Sunjevaric I, Alvaro D, Ciccone S, Wagner M, Rothstein R. 2011. Selective ploidy ablation, a high-throughput plasmid transfer protocol, identifies new genes affecting topoisomerase I-induced DNA damage. Genome Research 21:477–486. doi: 10.1101/gr.109033.110