Computational and Structural Biotechnology Journal 17 (2019) 805–811

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Mini Review

Inferring Protein-Protein Interaction Networks From Mass Spectrometry-Based Proteomic Approaches: A Mini-Review

Kumar Yugandhar [a,b,1], Shagun Gupta [a,b,1], Haiyuan Yu [a,b,*]

[a] Department of Computational Biology, Cornell University, Ithaca, New York 14853, USA
[b] Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, New York 14853, USA

ARTICLE INFO

Article history:
Received 2 March 2019
Received in revised form 20 May 2019
Accepted 26 May 2019
Available online 20 June 2019

Keywords:
Protein-protein
Interaction network
Systems biology
Mass spectrometry
Proteomics

ABSTRACT

Studying protein-protein interaction networks provide key evidence for the underlying molecular mechanisms. Mass spectrometry-based proteomic approaches have been playing a pivotal role in deciphering these interaction networks, along with precise quantification for individual interactions. In this mini-review we discuss the available techniques and methods for qualitative and quantitative elucidation of protein-protein interaction networks. We then summarize the down-stream computational strategies for identification and quantification of interactions from those techniques. Finally, we highlight the challenges and limitations of current computational pipelines in eliminating false positive interactors, followed by a summary of the innovative algorithms to address these issues, along with the scope for future improvements.

Contents

* Corresponding author at: Department of Computational Biology, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, New York 14853, USA
E-mail address: haiyuan.yu@cornell.edu (H. Yu).

[1] The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

https://doi.org/10.1016/j.csbj.2019.05.007

1. Introduction

Protein-protein interactions are intrinsic for fundamental cellular mechanisms. One of the primary goals of systems biology is to understand the functions of proteins from various organisms [1]. Some of the most widely used techniques to identify protein-protein interactions include yeast two-hybrid (Y2H) [2,3], protein-fragment complementation assay (PCA) [4], LUMIER [5], fluorescence resonance energy transfer (FRET) [6] etc. Mass spectrometry is a powerful tool for studying biomolecules such as proteins by their identification, quantification and further functional characterization [7–9]. The past two decades have witnessed a rapid progress in mass spectrometry-based methods for protein-protein interaction detection.

Mass spectrometry-based proteomics approaches can be broadly classified into two groups namely qualitative and quantitative. A conventional qualitative technique that has been used to study functions and interactions of proteins is Affinity purification mass spectrometry (AP-MS). AP-MS facilitates the isolation of protein complexes from cell lysates, which enables studying them at near physiological conditions [10]. AP-MS can also be used to study the dynamics of protein interactions, when combined with quantification approaches. Cross-linking mass spectrometry (XL-MS) is a more advanced technique to elucidate protein interaction networks at the level of both single complex and proteome-wide studies (extensively reviewed by Sinz [11]). Significant attention has been given towards designing MS techniques that would enable us to quantify the identified interactions. Such methods would provide invaluable information especially when two different biological conditions are being compared (e.g. normal vs disease states). Some of the most widely used quantification methods employ stable isotope labeling of amino acids in cell culture (SILAC) [12], isobaric tags for relative and absolute quantitation (iTRAQ) [13], tandem mass tag (TMT) [14] and protein correlation profiling (PCP) [15]. More recently, efforts are being made to design and utilize isotope-labelled cross-linkers to perform quantitative XL-MS [16–19].

In this mini-review we summarize some of the most widely used MS-based proteomic approaches to decipher protein-protein interaction networks and study the interaction dynamics. We further discuss various strategies that are being employed in the down-stream computational pipelines, along with potential challenges and the scope for improvements.

While the qualitative approaches provide information about the occurrence of a given interaction, quantitative approaches provide relative quantification of the interaction across multiple conditions (Fig. 1). Here, we discuss some of the most commonly used qualitative and labelled quantification approaches.

2. Qualitative Approaches

In this section, we discuss a traditional and classical technique known as AP-MS, which has been playing fundamental role in identifying protein interactions, along with a more advanced and rapidly emerging technique known as XL-MS.

3. AP-MS

AP-MS facilitates the identification of interactors for a protein of interest through affinity-based approaches [10,20]. Typical AP-MS workflow involves isolation of the protein of interest along with bound interactors using an affinity tag, followed by mass spectrometry to generate a list of potential interactors [10]. Further, in order to prioritize and segregate true positives from false positive interactors, various computational algorithms have been reported in the literature. More specifically, Gavin et al. [21] developed a 'socio-affinity' score to cluster the identified proteins into potential functional complexes and compared them with a manually curated set of known interactions to infer accuracy and coverage of the identified interactors. Motivated by the framework of socio-affinity score, Collins et al. [22] designed 'purification enrichment score' by incorporating features such as the negative evidence against interactions where one of the two proteins fail to be identified as a prey when another is used as a bait, and the probability of observing a pair of proteins in the same purification if those two proteins do not interact. On the other hand, constrained randomized simulations have also been demonstrated to be effective in generating co-occurrence distributions for each protein pair in a reference-set independent manner [23]. Furthermore, features such as the topological relationships between direct physical interactions along with observed co-complex interactions [24] and gene ontology enrichment [25] have also been employed to prioritize potential true positive interactors. More recently, an ensemble approach has been reported that utilizes multiple existing scoring methods

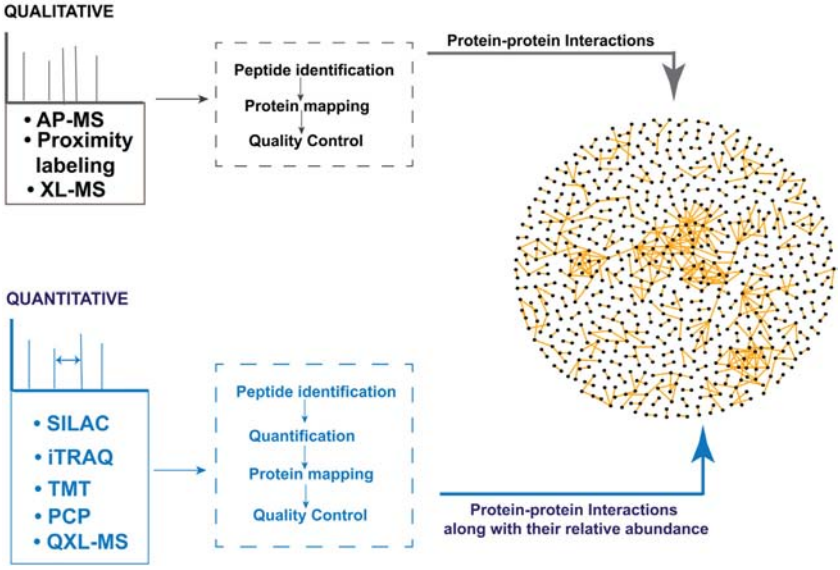

Fig. 1. Overview of current MS-based qualitative and quantitative proteomic approaches for elucidating interaction networks (AP-MS: affinity purification mass spectrometry; XL-MS: cross-linking mass spectrometry; SILAC: stable isotope labeling of amino acids in cell culture; iTRAQ: isobaric tags for relative and absolute quantitation; TMT: tandem mass tag; PCP: protein correlation profiling; QXL-MS: quantitative XL-MS).

to generate an initial network and then refine it further by applying indirect association removal methods [26].

4. Proximity Labeling

Despite being a highly popular approach, AP-MS suffers from limitations such as its inability to capture transient interactions and generation of potential non-specific (false positive) interactions between protein from multiple compartments during to the cell lysis step in sample preparation. These limitations can be partially addressed by in situ proximity labeling approaches such as BioID [27] and APEX [28]. Both BioID and APEX relies on a reactive biotin derivative that diffuses from an enzyme's active site to label proteins which are in spatial proximity [29]. However, the proximity labeling approaches come with their own caveats. Most importantly, since all the proteins that are in the vicinity of bait would be labelled and inferred as potential interactors, those labelled candidates need to be thoroughly evaluated by stringent downstream computational pipelines to eliminate false positive identifications.

5. XL-MS

In order to address the limited ability of AP-MS in capturing the weak or transient interactions, cross-linking approaches have been developed [10]. XL-MS typically utilizes a bi-functional reagent that covalently links two residues with reactive functional groups (commonly the primary amines of Lysine residues) that are within accessible distance. The cross-linked peptides are further identified using mass spectrometry to infer physical interactions and the structural constraints [30–32]. Efforts have been made to utilize cross-linking either as an additional step in the original protocol to covalently cross-link all the interactors to the bait protein [33] or as an independent method to identify proteome-wide interaction networks [34–38]. Typical analysis pipeline for mass spectrometry data analysis involves a database search to match the experimental spectra to the theoretical spectra for potential peptides from a protein database. This step becomes more challenging in case of XL-MS where two peptides need to be identified from a single spectrum, thereby increasing the probability for false positive identifications. Additionally, the database search needs to perform 2^n iterations (where n is the total number of peptides from a database), making it virtually impossible for proteome-wide studies. Later, the inception of MS-cleavable cross-linkers such as disuccinimidyl sulfoxide (DSSO) [39], disuccinimidyl dibutyric urea (DSBU) [40,41] and 1,1′-carbonyldiimidazole (CDI) [42], which generate a fragment ion signature that provides additional information for database search and reduce the number of iterations drastically (2^n to 2n), to facilitate proteomic XL-MS studies. Additionally, XL-MS has been demonstrated to be an efficient method to capture the distance constraints, thereby revealing crucial information that enables the identification of the interaction partners and dynamics of protein-protein interactions [43,44]. However, there is a great scope for improvement in the computational algorithms that are used to identify the cross-linked peptides especially by implementing novel and highly stringent error estimates [38,45–48].

6. Quantitative Approaches

Features such as spectral counts or integrated peptide ion intensities have been successfully utilized to infer quantitative information about the interactions identified through AP-MS experiments known as label-free quantification, using robust tools such as SAINT [49], MaxQuant [50], Census [51] etc. Furthermore, several labelled quantification approaches have been developed, which will be discussed in detail in the following sections.

7. SILAC

Stable Isotope Labeling by/with Amino acids in Cell culture (SILAC) is a popular method in quantitative proteomics that expanded on existing AP-MS, to provide more accurate identification of true protein interactors and contaminants. SILAC essentially allows artificial labelling of the peptides with Arginine and/or Lysine amino acid enriched media. This labelling with amino acids in comparison to traditional labelling with heavy carbon and nitrogen isotopes allows more efficient data interpretation through MS scans since the mass difference in the samples is unique and predictable, and hence has a distinctive isotopic distribution. By comparison of cells grown in “light” or control culture media (composed of arginine-0 and lysine-0), combination of arginine and lysine (for IP of interest) and of wild type and mutant IP in 1:1:1 ratios, SILAC protocol effectively deals with bias introduced by machine and human error [52].

Triple SILAC (SILAC with three isotope labelling states), Spike-in SILAC and Super-SILAC are some commonly used variants of the SILAC protocol. Triple labelling in SILAC can be performed under the conditions of pull-down without bait, pull-down with bait and pull-down with bait along with stimulus. This has been shown to be useful in identifying protein interactions that are stimulus-dependent [53]. It can also be used in investigating temporal proteomic changes, where the third cellular state is the time point of the treatment under study [54–56]. One of the earlier limitations of SILAC was that it could only be used to process cultured cells to allow complete incorporation of the isotope and couldn't be used with human tissue samples [57]. This has been addressed by the onset of Spike-in SILAC (which uses a stable isotope labelled cell line as a standard for comparison to in vitro post-isolation labelling of the cell line) [58] and Super-SILAC (which is run on a mixture of different cell lines and has the advantage over standard SILAC in not being limited by the number of samples that can be analyzed simultaneously) [57,59].

MaxQuant [60] is one of the most widely used tools for processing large-scale SILAC datasets [61], along with Census [51,62], pQuant [63], TPP [64], IsoQuant [65] etc. MaxQuant utilizes Andromeda search engine that incorporates length, charge and number of modifications of the peptides for quality peptide spectrum match (PSM) scoring [66]. Census and pQuant have functionalities that allow them to process ^{15}N labelled data. Furthermore, TPP supports both ETD (electron transfer dissociation) and CID (collision-induced dissociation) type of tandem MS data [67]. A significant limitation of SILAC methodology has been incomplete incorporation of the isotopes [67]. This has been addressed by approaches such as dataset normalization [68] and label-swap replication [69,70].

8. TMT and iTRAQ

A potential disadvantage with amino-group isotopic-labelling strategies is the increased sample complexity due to the double peaks in MS spectra [71,72]. This in turn magnifies the problem of undersampling. Isobaric tagging strategies such as Tandem Mass Tag (TMT) and Isobaric Tags for Relative and Absolute Quantitation (iTRAQ), which employ tags of equal mass, could overcome such problems [14,71]. Moreover with equal mass, the tags act as more efficient reciprocal standards (“light” vs “heavy”) for each other, especially since they elute together [14]. TMT belongs to a class of isobaric tags that allow for more accurate multiplexed quantification of IPs using MS^2 spectra [14,73,74]. Since the tags used in the methodology are isobaric and chemically equivalent, identical peptides labelled with different versions of the tag elute together during chromatography and can be cleaved from the peptide with collision-induced dissociation (CID) in tandem MS. Since quantification can be performed at the MS^2 level, TMT reduces the noise to signal ratios compared to the quantification at typically noisier MS^1. TMT is an improvement on previously available sampling techniques such as isotope coded affinity tag (ICAT). iTRAQ is indistinguishable from TMT

in terms of the precision of matching peptides to spectrum and they both deal with the problem of proteome dynamic range with comparable efficiency [75–77]. While TMT can be used for parallel multiplexing of up to 11 samples, iTRAQ can be used for up to 8 samples. However, some of the previous studies reported that higher order multiplexing such as iTRAQ 8-plex and TMT 6-plex yielded consistently lower number of protein and peptide identifications when compared to that of lower order multiplexing (iTRAQ 4-plex) [78,79].

Some of the key issues such as interference, where contaminants that elute with the target ion, can potentially skew the reporter ion intensities and contribute to underestimation of protein fold change estimations [71,80], in addition to the problem of ratio compression due to co-eluting peptides [81]. This problem could be addressed by performing an additional step of isolation and fragmentation (MS3 approaches) [82]. The key advantage of MS3 approaches is the minimal probability for the contaminating isobaric peptides to fragment into ions of the same mass as the target ion at MS^3 level [80]. Another potential issue is the variation in protein abundance, which makes it difficult to detect proteins with low abundance [80]. To address this problem, VSN (variance stabilizing normalization) has been proposed with functionalities such as outlier removal, log transformation, weighted means [83–86]. However it can aggravate the problem of underestimation of fold change of proteins [80]. Data modeling with 'boosted median' by Breitwieser et al. [87] and simple linear correction proposed by Karp et al. [88] have also been used for handling the problem of underestimation of protein fold change and high variation in protein abundance in samples.

9. Protein Correlation Profiling (PCP)

MS-PCP follows a co-fractionation based clustering approach to predict components in a specific protein complex, assuming protein complexes co-purify when separated based on their physicochemical properties such as density, hydrophobicity and size [89]. PCP was initially utilized by the Matthias Mann group to characterize the human centrosome [15]. Since then, PCP has been demonstrated to be an efficient tool in elucidation and characterization of protein complexes at both organelle [90] and cellular level [91], facilitating comparisons across multiple species [92]. Furthermore, PCP has also been employed in conjunction labelled approaches such as SILAC (PCP-SILAC) [93]. The key component in maintaining high specificity in a MS-PCP based approach is its integrated computational analysis pipeline employed for hierarchical clustering and the expression neighborhood analysis [90]. Typical computational workflow involves assignment of correlation scores for co-eluting proteins and utilize machine-learning based algorithms to perform clustering and annotate potential members of a predicted co-complex [92].

10. Quantitative XL-MS (QXL-MS)

QXL-MS is a relatively new technology, which utilizes labelled cross-linkers to quantify cross-linked peptides there by quantifying the protein interactions and the interaction dynamics [16]. Recent reports suggest that QXL-MS shows promising performance in studying the dynamics of human complement protein C3 [17–19]. However, the most challenging aspect for the downstream computational pipelines would be the typical low intensity of the cross-link peaks in the quantification spectra.

11. Specific Strategies and Challenges in Downstream Computational Pipelines

Morris et al. segregated the expected type of interactions from the MS-based protein interaction studies into four categories [94]. These categories include (i) the true interactions which, as the name suggests, are the interactions expected from the complex protein mixture under study as result of direct interaction between the protein with the bait (binary) or indirect interaction (co-complex) as a result of multiple proteins coming together in the cell to carry out specific functions, (ii) interactions as a result of carry-over contamination which is caused by the proteins/peptides from a previous experiment on the mass spectrometer [95], (iii) interactions due to general contaminants which are artifacts, caused by the background or environmental conditions in which the experiment is run and need to be eliminated from the set of observed interactions, and (iv) false positive interactions as a result of biases from mainly the degenerate peptides and stoichiometric variance. One of the biggest challenges for the computational pipelines is the reliable classification of the vast number of potential interactions from an MS experiment into these four categories. The following sections review the efforts and limitations in dealing with the carry-over contamination, general contaminants, and the false positive interactions. Furthermore, computational challenges, and the methodologies that address those challenges are summarized in Table 1, along with specific key features and the advantages.

12. Carry-Over Contamination

These contaminants are carried over from a previous experiment performed on the same mass spectrometer and typically emerge from experiments involving overexpression of proteins and involving hydrophobic proteins [94]. The obvious experimental strategies to address such contamination is to perform rigorous washing and sanitization of machines after each experiment [95], apart from regularly monitoring the liquid chromatography system [96]. Furthermore, it could potentially be addressed using computational approach by sorting the MS runs in the order they were run on the machine and delineate the signal intensities, number of spectra and number of identified unique peptides from proteins over consecutive MS runs and utilize that information as a predictive feature for the contamination [94,97]. Another method that has been proposed to assess and reduce carryover contamination involve introducing a "dummy ion transition" scan step while using old mass spectrometers where collision cells do not empty fast enough and lead to "cross talk" crossover contamination which can become problematic in case of shared fragment ions across different analytes [95,96,98].

13. Background Contamination

A subset of the interactors inferred from a typical MS-based proteomic experiment might include artifacts that could potentially result

Table 1
Summary of computational approaches addressing challenges in inferring interactions from MS-based proteomic data.

Challenge	Computational pipeline	Unique features/Advantages
Background contaminants	CRAPome [103]	List of proteins aggregated from negative control AP-MS experiments.
Carry-over contamination	Scanning for half-life-like patterns [94]	Analysis by identification of decreasing intensity, spectral counts, unique peptides over consecutive MS runs
Degenerate peptides	ProteinProphet [105]	Removes low PSM scoring spectra and calculates peptide scores from remaining spectra
	Scaffold [107]	"Greedy" approach
"One-hit" wonders	ProteinProphet [105]	Employs a probability-based model
	IDPicker [109]	Performs better than both "two-peptide" and "one peptide" rule
	HIquant [111]	Bottom-up approach; requires corresponding ion ratios
	MRM [115]	Requires prior knowledge of all protein stoichiometries/isoforms
	SWATH MS [116]	Targeted MS
Saturation of spectral peaks	SignalFinder DeconTools [117]	Non-targeted MS

from non-specific binding during the affinity purification step [99]. This might include non-specific binding of proteins to beads, resins or might result from biological factors such as protein misfolding [94]. Additionally, non-specific interactions between proteins from different cellular compartments could be induced by the lysis step in a typical AP-MS experimental procedure. Multiple experimental strategies have been developed to distinguish specific and non-specific interactions to filter-out the background contaminants [100–102]. Few years ago, Mellacheruvu et al. [103] published a comprehensive list of background proteins, called CRAPome, consisting artifacts of human or biological errors, aggregated from multiple AP-MS experiments. This resource allows users to query a protein of interest or download the list of contaminants for selected experimental conditions and perform data analysis.

14. False Positive Interactions

It is of paramount importance to identify and eliminate false positive interactions, such that the interaction networks generated using these data would facilitate reliable computational analyses. The following subsections briefly describe potential sources that result in false positives, such as (a) degenerate peptides (b) "one-hit" wonders, (c) stoichiometric variance, and (d) saturation of the spectral peaks.

14.1. Degenerate Peptides

Degenerate peptides are the peptides that are shared by multiple proteins and their presence implicates the presence of all the proteins that share these peptides [104]. Multiple approaches have been reported to address the issue of degenerate peptides. Algorithms such as ProteinProphet [105] for example, deals with degenerate peptides by retaining only the spectrum associated with the highest PSM scores and uses the remaining spectra to calculate scores for peptides as an approximation of their probabilities [106]. Another tool called Scaffold [107] address this issue using a "greedy" approach, wherein protein scores are first calculated using peptides that do not fall under the category of degenerate peptides. Then peptides belonging to the class of degenerate peptides are deterministically assigned to the protein that scores the highest out of all the proteins that share the peptide [106].

14.2. "One-Hit" Wonders

Another challenge for computational pipelines in distinguishing true interactors from false positives is the case where multiple potential interactor proteins have only one peptide (unique or shared) identified from the PSM search [104]. A general approach that has been adapted to deal with such cases is the 'two peptide' rule, where only proteins with two or more number of identified peptides are retained [108]. ProteinProphet employs more complex statistical and machine learning approaches to alleviate the issue of false positives [105]. IDPicker is another pipeline that has been reported to show better specificity compared to the "one peptide" and "two peptide" rules, however showing lower sensitivity than both those approaches [106,109].

14.3. Stoichiometric Variance

Complex stoichiometry is a quantitative measure that refers to the relative amount of the constituents of a complex protein mixture [110]. The major challenge for computational pipelines in analyzing proteomic data is the identification and determination of the stoichiometry of protein complexes that are composed of protein isoforms, especially since the protein isoforms can have distinct biological functions [111]. This issue is especially hard to deal in a bottom-up approach, where peptides generated due to enzymatic digestion could be shared between multiple proteoforms [111]. Methods such as HIquant [111] have been developed that can be used to differentiate protein isoforms, identifiable by very few unique peptides using information about ratios between chemically identical ions. Additionally, tools such as "multiple reaction monitoring" (MRM) and "sequential window acquisition of all theoretical fragment-ion spectra" (SWATH MS) have also been proposed as effective strategies for determining proteoforms from a complex mixture [112].

14.4. Saturation of the Spectral Peaks That Might Lead to Under/Over Estimation

Occasionally, the light or heavy peak signal in the quantification spectra could exceed the signal threshold of the detector for species in high concentration, leading to saturated peaks and disposal of useful hits, which in turn might lead to erroneous identification and quantification [113]. However, it is important to note that this problem might not be applicable for majority of the currently available advanced mass spectrometers. In order to address the issue, algorithms such as SignalFinder have been implemented for targeted MS runs. On the other hand, tools used for non-targeted MS runs generally involve usage of adjacent ion peaks that are not saturated to correct the saturated peaks [113,114]. Furthermore, software such as DeconTools correct saturated peaks above a pre-determined threshold by utilising both the unsaturated peaks and theoretical isotopic peaks of the peptides [113].

15. Conclusion

In this mini-review we discussed some of the most widely used mass spectrometry-based approaches for inferring protein-protein interaction networks. With the rapid advancement of instrument technology, the mass spectrometers are evolving with exciting and powerful capabilities, thus greatly increasing the sensitivity of proteomic studies. However, the higher sensitivity of proteomics data sets poses newer challenges for the downstream computational pipelines, most importantly detecting and eliminating contaminants and false positives. We reviewed some of the key challenges that are general to MS-based proteomic approaches, along with difficulties in analyzing data from specific type of experiments. Furthermore, we discussed the efforts and innovations that have been developed and adapted by the scientific community to address those problems. While we currently witness several novel technological and methodological advancements in the field, there is great scope for further improvement especially for the computational analytical pipelines. Finally, we propose effective utilization of the currently known binary and co-complex interactions for systematic validation of large-scale interaction studies, which would in turn contribute to the growth of high-quality interaction datasets, facilitating reliable network analyses.

Declarations of Competing Interest

None.

Acknowledgements

KY thanks the Sam and Nancy Fleming Research Fellowship. National Institutes of Health grants (R01 GM097358, R01 GM104424, R01 GM124559, R01 GM125639, R01 DK115398, R01 CA167824, R01 HD082568, UM1 HG009393); National Science Foundation grant (DBI-1661380) and Simons Foundation Autism Research Initiative grant (575547) to HY

References

[1] Yu H, et al. High-quality binary protein interaction map of the yeast interactome network. Science 2008;322:104.
[2] Fields S, Song, O.-k.. A novel genetic system to detect protein–protein interactions. Nature 1989;340:245–6.

[3] Chien CT, Bartel PL, Sternglanz R, Fields S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc Natl Acad Sci 1991;88:9578.
[4] Michnick SW, Remy I, Campbell-Valois F-X, Vallée-Bélisle A, Pelletier JN. In: Thorner J, Emr SD, Abelson JN, editors. Methods in enzymology. Academic Press; 2000. p. 208–30.
[5] Barrios-Rodiles M, et al. High-throughput mapping of a dynamic signaling network in mammalian cells. Science 2005;307:1621.
[6] Selvin PR. Methods in enzymology. , vol. 246Academic Press; 1995; 300–34.
[7] Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 2003;422:198.
[8] Savaryn JP, Toby TK, Kelleher NL. A researcher's guide to mass spectrometry-based proteomics. Proteomics 2016;16:2435–43.
[9] Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. Nature 2016;537:347.
[10] Gingras AC, Gstaiger M, Raught B, Aebersold R. Analysis of protein complexes using mass spectrometry. Nat Rev Mol Cell Biol 2007;8:645–54.
[11] Sinz A. Cross-linking/mass spectrometry for studying protein structures and protein–protein interactions: where are we now and where should we go from Here? Angew Chem Int Ed 2018;57:6390–6.
[12] Ong S-E, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 2002;1:376.
[13] Wiese S, Reidegeld KA, Meyer HE, Warscheid B. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. Proteomics 2007;7:340–50.
[14] Thompson A, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 2003;75:1895–904.
[15] Andersen JS, et al. Proteomic characterization of the human centrosome by protein correlation profiling. Nature 2003;426:570–4.
[16] Fischer L, Chen ZA, Rappsilber J. Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. J Proteomics 2013;88:120–8.
[17] Chen Z, et al. Quantitative cross-linking/mass spectrometry reveals subtle protein conformational changes. Well Open Res 2016;1:5.
[18] Chen ZA, Rappsilber J. Protein dynamics in solution by quantitative crosslinking/mass spectrometry. Trends Biochem Sci 2018;43:908–20.
[19] Chen ZA, Rappsilber J. Quantitative cross-linking/mass spectrometry to elucidate structural changes in proteins and their complexes. Nat Protoc 2019;14:171–201.
[20] Dunham WH, Mullin M, Gingras A-C. Affinity-purification coupled to mass spectrometry: basic principles and strategies. PROTEOMICS 2012;12:1576–90.
[21] Gavin A-C, et al. Proteome survey reveals modularity of the yeast cell machinery. Nature 2006;440:631.
[22] Collins SR, et al. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. Mol Cell Proteomics 2007;6:439.
[23] Yu X, Ivanic J, Wallqvist A, Reifman J. A novel scoring approach for protein co-purification data reveals high interaction specificity. PLoS Comput Biol 2009;5:e1000515.
[24] Zhang X-F, Ou-Yang L, Hu X, Dai D-Q. Identifying binary protein-protein interactions from affinity purification mass spectrometry data. BMC Genomics 2015;16:745.
[25] Zhang B, Park B-H, Samatova NF, Karpinets T. From pull-down data to protein interaction networks and complexes with biological relevance. Bioinformatics 2008;24:979–86.
[26] Tian B, Zhao C, Gu F, He Z. A two-step framework for inferring direct protein-protein interaction network from AP-MS data. BMC Syst Biol 2017;11:82.
[27] Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol 2012;196:801.
[28] Rhee H-W, et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 2013;339:1328.
[29] Trinkle-Mulcahy L. Recent advances in proximity-based labeling methods for interactome mapping. F1000Research 2019;8 F1000 Faculty Rev-1135.
[30] Rappsilber J. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J Struct Biol 2011;173:530–40.
[31] Sinz A, Arlt C, Chorev D, Sharon M. Chemical cross-linking and native mass spectrometry: a fruitful combination for structural biology. Protein Sci 2015;24:1193–209.
[32] Barysz HM, Malmström J. Development of large-scale cross-linking mass spectrometry. Mol Cell Proteomics 2018;17:1055.
[33] Makowski MM, Willems E, Jansen PWTC, Vermeulen M. Cross-linking immunoprecipitation-MS (xIP-MS): topological analysis of chromatin-associated protein complexes using single affinity purification. Mol Cell Proteomics 2016;15:854.
[34] Navare Arti T, et al. Probing the protein interaction network of *Pseudomonas aeruginosa* cells by chemical cross-linking mass spectrometry. Structure 2015;23:762–73.
[35] Liu F, Rijkers DTS, Post H, Heck AJR. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. Nat Methods 2015;12:1179.
[36] Liu F, Lössl P, Scheltema R, Viner R, Heck AJR. Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. Nat Commun 2017;8:15473.
[37] Weisbrod CR, et al. In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. J Proteome Res 2013;12:1569–79.
[38] Yugandhar K, et al. MaXLinker: proteome-wide cross-link identifications with high specificity and sensitivity. bioRxiv 2019:526897.
[39] Kao A, et al. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. Mol Cell Proteomics 2011;10 M110.002212.
[40] Müller MQ, Dreiocker F, Ihling CH, Schäfer M, Sinz A. Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS. Anal Chem 2010;82:6958–68.
[41] Müller MQ, et al. A universal matrix-assisted laser desorption/ionization cleavable cross-linker for protein structure analysis. Rapid Commun Mass Spectrom 2011;25:155–61.
[42] Hage C, Iacobucci C, Rehkamp A, Arlt C, Sinz A. The first zero-length mass spectrometry-cleavable cross-linker for protein structure analysis. Angew Chem Int Ed 2017;56:14551–5.
[43] Sinz A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. Mass Spectrom Rev 2006;25:663–82.
[44] O'Reilly FJ, Rappsilber J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. Nat Struct Mol Biol 2018;25:1000–8.
[45] Walzthoeni T, et al. False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat Methods 2012;9:901.
[46] Fischer L, Rappsilber J. Quirks of error estimation in cross-linking/mass spectrometry. Anal Chem 2017;89:3829–33.
[47] Trnka MJ, Baker PR, Robinson PJJ, Burlingame AL, Chalkley RJ. Matching cross-linked peptide spectra: only as good as the worse identification. Mol Cell Proteomics 2014;13:420.
[48] Iacobucci C, Sinz A. To be or not to be? Five guidelines to avoid misassignments in cross-linking/mass spectrometry. Anal Chem 2017;89:7832–5.
[49] Choi H, et al. Analyzing protein-protein interactions from affinity purification-mass spectrometry data with SAINT. Current protocols in bioinformatics; 2012 Chapter 8, Unit8.15-Unit18.15.
[50] Cox J, et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics 2014;13:2513–26.
[51] Park SK, Venable JD, Xu T, Yates 3rd JR. A quantitative analysis software tool for mass spectrometry-based proteomics. Nat Methods 2008;5:319–22.
[52] ten Have S, Boulon S, Ahmad Y, Lamond AI. Mass spectrometry-based immunoprecipitation proteomics - the user's guide. Proteomics 2011;11:1153–9.
[53] Hilger M, Mann M. Triple SILAC to determine stimulus specific interactions in the Wnt pathway. J Proteome Res 2012;11:982–94.
[54] Lanucara F, Eyers CE. In: Jameson D, Verma M, Westerhoff HV, editors. Methods in enzymology. Academic Press; 2011. p. 133–50.
[55] Blagoev B, Ong S-E, Kratchmarova I, Mann M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol 2004;22:1139.
[56] Wang X, et al. SILAC-based quantitative MS approach for real-time recording protein-mediated cell-cell interactions. Sci Rep 2018;8:8441.
[57] Pozniak Y, Geiger T. In: Warscheid B, editor. Stable isotope labeling by amino acids in cell culture (SILAC): methods and protocols. New York, NY: Springer New York; 2014. p. 281–91.
[58] Pinho JPC, Bell-Temin H, Liu B, Stevens SM. In: Kobeissy FH, Stevens JSM, editors. Neuroproteomics: methods and protocols. New York, NY: Springer New York; 2017. p. 295–312.
[59] Shenoy A, Geiger T. Super-SILAC: current trends and future perspectives. Expert Rev Proteomics 2015;12:13–9.
[60] Tyanova S, Mann M, Cox J. In: Warscheid B, editor. Stable isotope labeling by amino acids in cell culture (SILAC): methods and protocols. New York, NY: Springer New York; 2014. p. 351–64.
[61] Deng J, Erdjument-Bromage H, Neubert TA. Quantitative comparison of proteomes using SILAC. Curr Protoc Protein Sci 2019;95:e74.
[62] Park SK, Yates Iii JR. Census for proteome quantification. Curr Protoc Bioinformatics 2010;29:13.12.11.
[63] Liu C, et al. pQuant improves quantitation by keeping out interfering signals and evaluating the accuracy of calculated ratios. Anal Chem 2014;86:5286–94.
[64] Keller A, Shteynberg D. In: Wu CH, Chen C, editors. Bioinformatics for comparative proteomics. Totowa, NJ: Humana Press; 2011. p. 169–89.
[65] Liao Z, Wan Y, Thomas SN, Yang AJ. IsoQuant: a software tool for stable isotope labeling by amino acids in cell culture-based mass spectrometry quantitation. Anal Chem 2012;84:4535–43.
[66] Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 2016;11:2301.
[67] Chen X, Wei S, Ji Y, Guo X, Yang F. Quantitative proteomics using SILAC: principles, applications, and developments. PROTEOMICS 2015;15:3175–92.
[68] Pasculescu A, et al. CoreFlow: a computational platform for integration, analysis and modeling of complex biological data. J Proteomics 2014;100:167–73.
[69] Schulze WX, Mann M. A novel proteomic screen for peptide-protein interactions. J Biol Chem 2004;279:10756–64.
[70] Park S-S, et al. Effective correction of experimental errors in quantitative proteomics using stable isotope labeling by amino acids in cell culture (SILAC). J Proteomics 2012;75:3720–32.
[71] Treumann A, Thiede B. Isobaric protein and peptide quantification: perspectives and issues. Expert Rev Proteomics 2010;7:647–53.
[72] Tate S, Larsen B, Bonner R, Gingras A-C. Label-free quantitative proteomics trends for protein–protein interactions. J Proteomics 2013;81:91–101.
[73] Zhang L, Elias JE. In: Comai L, Katz JE, Mallick P, editors. Proteomics: methods and protocols. New York, NY: Springer New York; 2017. p. 185–98.

[74] Lau H-T, Suh HW, Golkowski M, Ong S-E. Comparing SILAC- and stable isotope dimethyl-labeling approaches for quantitative proteomics. J Proteome Res 2014;13:4164–74.
[75] Rauniyar N, Yates 3rd JR. Isobaric labeling-based relative quantification in shotgun proteomics. J Proteome Res 2014;13:5293–309.
[76] Wu L, Han DK. Overcoming the dynamic range problem in mass spectrometry-based shotgun proteomics. Expert Rev Proteomics 2006;3:611–9.
[77] Zubarev RA. The challenge of the proteome dynamic range and its implications for in-depth proteomics. Proteomics 2013;13:723–6.
[78] Pichler P, et al. Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-Plex compared to TMT 6-Plex and iTRAQ 8-Plex on LTQ Orbitrap. Anal Chem 2010;82:6549–58.
[79] Casey TM, et al. Analysis of reproducibility of proteome coverage and quantitation using isobaric mass tags (iTRAQ and TMT). J Proteome Res 2017;16:384–92.
[80] Evans C, et al. An insight into iTRAQ: where do we stand now? Anal Bioanal Chem 2012;404:1011–27.
[81] Ow SY, Salim M, Noirel J, Evans C, Wright PC. Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation. PROTEOMICS 2011;11:2341–6.
[82] Ting L, Rad R, Gygi SP, Haas W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat Methods 2011;8:937–40.
[83] Boehm AM, Pütz S, Altenhöfer D, Sickmann A, Falk M. Precise protein quantification based on peptide quantification using iTRAQ™. BMC Bioinformatics 2007;8:214.
[84] Bantscheff M, et al. Robust and sensitive iTRAQ quantification on an LTQ orbitrap mass spectrometer. Mol Cell Proteomics 2008;7:1702.
[85] Choe LH, Aggarwal K, Franck Z, Lee KH. A comparison of the consistency of proteome quantitation using two-dimensional electrophoresis and shotgun isobaric tagging in Escherichia coli cells. Electrophoresis 2005;26:2437–49.
[86] Lin W-T, et al. Multi-Q: a fully automated tool for multiplexed protein quantitation. J Proteome Res 2006;5:2328–38.
[87] Breitwieser FP, et al. General statistical modeling of data from protein relative expression isobaric tags. J Proteome Res 2011;10:2758–66.
[88] Karp NA, et al. Addressing accuracy and precision issues in iTRAQ quantitation. Mol Cell Proteomics 2010;9:1885.
[89] Larance M, Lamond AI. Multidimensional proteomics for cell biology. Nat Rev Mol Cell Biol 2015;16:269.
[90] Foster LJ, et al. A mammalian organelle map by protein correlation profiling. Cell 2006;125:187–99.
[91] Havugimana Pierre C, et al. A census of human soluble protein complexes. Cell 2012;150:1068–81.
[92] Wan C, et al. Panorama of ancient metazoan macromolecular complexes. Nature 2015;525:339.
[93] Kristensen AR, Foster LJ. In: Warscheid B, editor. Stable isotope labeling by amino acids in cell culture (SILAC): methods and protocols. New York, NY: Springer New York; 2014. p. 263–70.
[94] Morris JH, et al. Affinity purification-mass spectrometry and network analysis to understand protein-protein interactions. Nat Protoc 2014;9:2539–54.
[95] Hughes NC, Wong EYK, Fan J, Bajaj N. Determination of carryover and contamination for mass spectrometry-based chromatographic assays. AAPS J 2007;9:E353–60.
[96] Bittremieux W, et al. Quality control in mass spectrometry-based proteomics. Mass Spectrom Rev 2018;37:697–711.
[97] Verschueren E, et al. Scoring large-scale affinity purification mass spectrometry datasets with MiST. Curr Protoc Bioinformatics 2015;49:8.19.11–18.19.16.
[98] Min SC, Kim EJ, El-Shourbagy TA. Managing bioanalytical cross-contamination, vol. 10; 2007.
[99] Chang I-F. Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes. PROTEOMICS 2006;6:6158–66.
[100] Selbach M, Mann M. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). Nat Methods 2006;3:981.
[101] Trinkle-Mulcahy L. Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry. PROTEOMICS 2012;12:1623–38.
[102] Tackett AJ, et al. I-DIRT, A general method for distinguishing between specific and nonspecific protein interactions. J Proteome Res 2005;4:1752–6.
[103] Mellacheruvu D, et al. The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. Nat Methods 2013;10:730.
[104] Huang T, Wang J, Yu W, He Z. Protein inference: a review. Brief Bioinform 2012;13:586–614.
[105] Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003;75:4646–58.
[106] Serang O, Noble W. A review of statistical methods for protein identification using tandem mass spectrometry. Stat Interface 2012;5:3–20.
[107] Searle BC. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. Proteomics 2010;10:1265–9.
[108] Gupta N, Pevzner PA. False discovery rates of protein identifications: a strike against the two-peptide rule. J Proteome Res 2009;8:4173–81.
[109] Zhang B, Chambers MC, Tabb DL. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. J Proteome Res 2007;6:3549–57.
[110] Wohlgemuth I, Lenz C, Urlaub H. Studying macromolecular complex stoichiometries by peptide-based mass spectrometry. Proteomics 2015;15:862–79.
[111] Malioutov D, et al. Quantifying homologous proteins and Proteoforms. Mol Cell Proteomics 2019;18:162–8.
[112] Stastna M, Van Eyk JE. Analysis of protein isoforms: can we do it better? Proteomics 2012;12:2937–48.
[113] Bilbao A, et al. An algorithm to correct saturated mass spectrometry ion abundances for enhanced quantitation and mass accuracy in omic studies. Int J Mass Spectrom 2018;427:91–9.
[114] Colangelo CM, Chung L, Bruce C, Cheung K-H. Review of software tools for design and analysis of large scale MRM proteomic datasets. Methods (San Diego, Calif) 2013;61:287–98.
[115] Kitteringham NR, Jenkins RE, Lane CS, Elliott VL, Park BK. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. J Chromatogr B 2009;877:1229–39.
[116] Gillet LC, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics 2012;11 O111.016717-O016111.016717.
[117] Jaitly N, et al. Decon2LS: an open-source software package for automated processing and visualization of high resolution mass spectrometry data. BMC Bioinformatics 2009;10:87.