
Faculty of Science

Faculty Publications

Interfaces with Structure Dynamics of the Workhorses from Cells Revealed through Cross-Linking Mass Spectrometry (CLMS)

Umesh Kalathiya, Monikaben Padariya, Jakub Faktor, Etienne Coyaoud, Javier A. Alfaro, Robin Fahraeus, ... & David R. Goodlett

March 2021

© 2021 Umesh Kalathiya et al. This is an open access article distributed under the terms of the Creative Commons Attribution License <https://creativecommons.org/licenses/by/4.0/>

This article was originally published at:




<https://doi.org/10.3390/biom11030382>

Citation for this paper:

Kalathiya, U., Padariya, M., Faktor, J., Coyaoud, E., Alfaro, J. A., Fahraeus, R., ... Goodlett, D. R. (2021). Interfaces with Structure Dynamics of the Workhorses from Cells Revealed through Cross-Linking Mass Spectrometry (CLMS). *Biomolecules*, 11(3), 1-27. <https://doi.org/10.3390/biom11030382>.

Review

Interfaces with Structure Dynamics of the Workhorses from Cells Revealed through Cross-Linking Mass Spectrometry (CLMS)

Umesh Kalathiya ^{1,*},[†] , Monikaben Padariya ^{1,†}, Jakub Faktor ¹ , Etienne Coyaoud ², Javier A. Alfaro ^{1,3}, Robin Fahraeus ¹, Ted R. Hupp ^{1,3}  and David R. Goodlett ^{1,4,5,*}

¹ International Centre for Cancer Vaccine Science, University of Gdansk, ul. Kładki 24, 80-822 Gdansk, Poland; monikaben.padariya@ug.edu.pl (M.P.); jakub.faktor@ug.edu.pl (J.F.); javier.alfaro@proteogenomics.ca (J.A.A.); robin.fahraeus@inserm.fr (R.F.); ted.hupp@ed.ac.uk (T.R.H.)

² Protéomique Réponse Inflammatoire Spectrométrie de Mass—PRISM, Inserm U1192, University Lille, CHU Lille, F-59000 Lille, France; etienne.coyaoud@inserm.fr

³ Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland EH4 2XR, UK

⁴ Department of Biochemistry & Microbiology, University of Victoria, Victoria, BC V8Z 7X8, Canada

⁵ Genome BC Proteome Centre, University of Victoria, Victoria, BC V8Z 5N3, Canada

* Correspondence: umesh.kalathiya@ug.edu.pl (U.K.); goodlett@uvic.ca (D.R.G.)

† These authors contributed equally to this work.



Citation: Kalathiya, U.; Padariya, M.; Faktor, J.; Coyaoud, E.; Alfaro, J.A.; Fahraeus, R.; Hupp, T.R.; Goodlett, D.R. Interfaces with Structure Dynamics of the Workhorses from Cells Revealed through Cross-Linking Mass Spectrometry (CLMS). *Biomolecules* **2021**, *11*, 382. <https://doi.org/10.3390/biom11030382>

Academic Editor: Vladimir N. Uversky

Received: 30 January 2021

Accepted: 1 March 2021

Published: 4 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The fundamentals of how protein–protein/RNA/DNA interactions influence the structures and functions of the workhorses from the cells have been well documented in the 20th century. A diverse set of methods exist to determine such interactions between different components, particularly, the mass spectrometry (MS) methods, with its advanced instrumentation, has become a significant approach to analyze a diverse range of biomolecules, as well as bring insights to their biomolecular processes. This review highlights the principal role of chemistry in MS-based structural proteomics approaches, with a particular focus on the chemical cross-linking of protein–protein/DNA/RNA complexes. In addition, we discuss different methods to prepare the cross-linked samples for MS analysis and tools to identify cross-linked peptides. Cross-linking mass spectrometry (CLMS) holds promise to identify interaction sites in larger and more complex biological systems. The typical CLMS workflow allows for the measurement of the proximity in three-dimensional space of amino acids, identifying proteins in direct contact with DNA or RNA, and it provides information on the folds of proteins as well as their topology in the complexes. Principal CLMS applications, its notable successes, as well as common pipelines that bridge proteomics, molecular biology, structural systems biology, and interactomics are outlined.

Keywords: cross-linking mass spectrometry; proteomics; chemical cross-linkers; CLMS; protein–protein; protein–DNA; protein–RNA interactions; structural biology

1. Introduction

Decades of research into the cell biology, molecular biology, biochemistry, structural biology, and biophysics have produced a detailed understanding of individual DNA/RNA/protein molecules, and their interconnected networks. A great diversity of techniques has emerged for studying their structural interactions. However, even more complex, the structures of these workhorses from the cells are themselves dynamic, converting from one dominant form to another based on the proportions of particular proteoforms present for any given biomolecules. Accordingly, beyond the protein–protein/DNA/RNA interaction landscape, there is an entire universe to explore with respect to their structure and dynamics. One such high-throughput technique has emerged as a dominant player in understanding both interaction landscapes and their resulting protein/DNA/RNA structures, namely cross-linking mass spectrometry (CLMS). This review covers different

methods that are available to study protein–protein/DNA/RNA interactions, and provides vital insight into CLMS, a collection of methods that are perfectly suited to achieve a better understanding of intra- or inter- molecular interactions.

Protein–protein interactions (PPIs) play a crucial role in all biological/biomolecular processes to understand the molecular mechanism of relevant protein molecules; hence, they are often termed the workhorses of cells. Over 80% of proteins do not function in isolation, but rather exist in interactions with one another to obtain stable or transitory complexes, as demanded by their observed function [1,2]. PPIs are considered to be an emerging class of drug target, since aberrant protein–protein interactions can participate in the pathogenesis of various human diseases, which, in turn, can contribute significant options for diagnostic as well as therapeutic targets. A large number of experimental methods have been developed to study PPIs, based on biophysical, biochemical, or genetic principles (as shown in Figure 1); and, each individual type has advantages as well as limitations regarding structural coverage of amino acid sequences, sensitivity, and specificity [1,3–6]. Methods can be chosen to emphasize different aspects of PPIs, such as identifying a protein binding partner(s), generating structural details of protein complexes, analyzing kinetic and thermodynamic constants of interactions, visualizing and quantifying PPIs in real time in living cells, and mapping small interactomes that refer to specific cellular pathways [1,3–5].

Mass spectrometry (MS) is one of the powerful approaches available that is useful in several areas beyond CLMS (e.g., hydrogen-deuterium exchange [7]), and it is particularly useful in combination with other techniques, providing steady progress for structural biology. Technical advances in mass spectrometry have made it possible to study protein–protein interactions from simple protein complexes to wide scale proteome experiments, which were not earlier accessible by traditional techniques [8–10]. Above all, mass spectrometry methods have democratized protein interaction analysis by making them accessible, relatively inexpensive, and high throughput. Having these advantages MS is becoming progressively popular in the structural biology stream for analyzing three dimensional (3D) structures and mapping their interactions with partner molecules. While emerging techniques, like cryoEM, can summarize large numbers of protein complexes, much of the resulting dynamics of the structure is missed and such gaps in structural data sets can be bridged by low-resolution methods, for example, chemical cross-linking (CL) [11–13]. The overall architecture of a protein complex can be obtained through electron microscopy (EM) [14], small-angle X-ray scattering (SAXS) [15], and ion-mobility (IM) MS [16], whereas the precise residues forming protein–protein interactions can be identified using hydrogen-deuterium exchange [17], chemical cross-linking [11–13], and chemical foot printing [18].

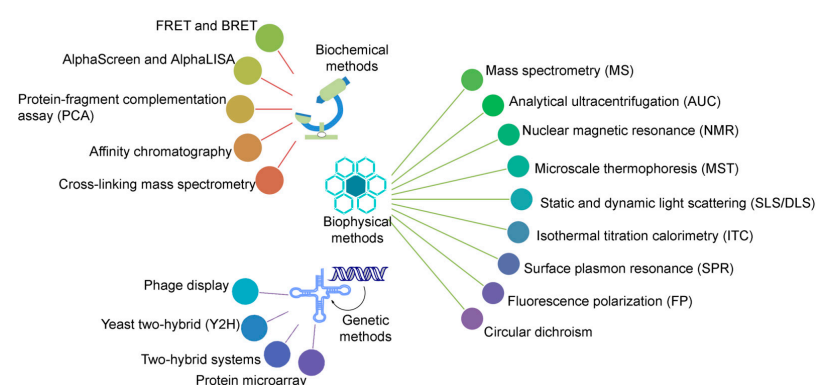


Figure 1. A spectrum of widely characterized experimental methods, based on biochemical, biophysical, or genetic principles. The listed methods define protein–protein interactions at various degrees of affinity as well as specificity [1,2,11–13,19–40]. The RNA and DNA diagrams, representing the genetic methods, were prepared using the BIOVIA draw (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) tool.

Fundamentally, the mass spectrometry methods simply measure a mass-to-charge ratio of an ion. Initially, analytes are ionized and then transferred into the gas phase prior to their separation according to mass-to-charge ratios in a mass analyzer. Subsequently, ions that emerge from the mass analyzer are recorded using a detector. The most common way to ionize a peptide or protein sample is electrospray ionization (ESI) [41]. However, a complementary ionization technique, known as matrix-assisted laser desorption/ionization (MALDI), is often used for its relative ease of use for novices. In a few studies analyzing intact protein complexes, MALDI mass spectrometry (MALDI-MS) has been used together with chemical cross-linking techniques [42–44]. Regardless of ionization methods, there are various types of mass analyzers accessible, including the following common ones: ion trap, time-of-flight (TOF), quadrupole, and orbitrap. Each of these mass analyzers may be configured in various ways with similar or different mass analyzers to form unique types of mass spectrometers [45].

Quantitative cross-linking mass spectrometry (QCLMS) approaches investigate protein structures as well as the dynamics of their interactions [46–51]. QCLMS is often performed using a unique cross-linker that introduces a corresponding mass shift after isotope labeling specific only to the cross-linked peptides [46,47,50,52,53], followed by quantitation of cross-links in MS1. However, the limited availability of isotope labeled cross-linkers restrains the implementation of this approach in QCLMS [54]. Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC), which is a popular option is an alternative to the QCLMS approach, which relies on the metabolic incorporation of isotope labeled amino acids from culture media. Metabolically labeled samples are cross-linked and pooled; relative quantitation is performed from MS1 data using the characteristic mass shift introduced into peptides from the incorporation of isotope labeled amino acid/s [48]. SILAC enables a comparison of multiple samples per analysis (usually two) and it can also enable monitoring amino acid incorporation in a time course (pulsed SILAC), which is especially valuable for exploring dynamics in biological processes. Recently, iCLASPI (in vivo cross-linking assisted and stable isotope labeling by amino acids in cell culture (SILAC)-based protein identification), an approach combining SILAC and in vivo cross-linking, has been implemented to quantify native protein–protein interactions in HEK293T cells. iCLASPI has been successfully implemented to profile native protein–protein interactions involving core histones H3 and H4 in biological context [55]. Chavez et al. successfully implemented SILAC and cross-linking to investigate key protein–protein interactions, and then investigated the Hsp90 conformational changes upon treatment with 17-AAG Hsp90 N-terminal domain inhibitor [49].

Isobaric Tags for Relative and Absolute Quantitation/Tandem Mass Tag (iTRAQ/TMT) are chemical labeling, which are introduced to peptides after protease digestion, which allow relative protein quantitation via MS2 or MS3 encoded data. Nowadays, TMT/iTRAQ labeling enables comparison of up to 16 samples (valid for TMT) in a single MS analysis. Notably, the iTRAQ/TMT reporter ions used for quantitation are cleaved from labeled peptides during fragmentation by way of collision-induced dissociation allowing for the quantitation from fragment ion relative intensities. Yu et al. [56], implemented the TMT approach in a multiplexed comparison of protein complex dynamics and protein–protein interactions. Their QMIX (Quantitation of Multiplexed, Isobaric-labeled cross (X)-linked peptides) workflow with TMT labeling, achieves peptide quantitation from MS3 data that eliminates interference from ions that were observed in MS1 data along with isotope labeled cross-linkers or SILAC cross-linking [56]. Furthermore, the precise MS2 quantitation of protein cross-linking could also be achieved in a label-free manner utilizing a data-independent acquisition mode (DIA). The extraction of cross-linked peptide quantities from DIA data is usually performed utilizing a spectral library prepared particularly from investigated samples. Muller et al. developed a novel DIA-QCLMS approach utilizing photo activatable cross-linkers ensuring reliable quantitation of cross-linked proteins across a wide range of environmental changes, such as pH, temperature pressure, or concentration [51,52,57].

DIA approaches promise to be useful in future applications of quantitative cross-linking proteomics, due to their precision, reproducibility, and label-free manner.

Moreover, techniques, such as proximity-dependent biotin labeling (or BioID technique) in living cells, help to understand the plasticity of protein networks within heterogeneous cellular populations. In combination with nanopore technology, such an approach could help tackle pending biological questions, e.g., the identification of peptidyl-prolyl isomerases (PPIases) substrates. PPIases substrates preserve their primary structure/molecular mass, as well as the *cis* and *trans* isomers of the proline peptide bonds of the substrates interconversion by PPIases being the sole change. This subtle modification triggers important changes to the substrate's fate, such as subcellular translocation, degradation, or rewiring of their protein–protein interaction networks. PPIase enzymes act as central molecular switches, as exemplified by the peptidyl-prolyl *cis-trans* isomerase NIMA-interacting 1 (Pin1) that has been extensively studied and showing its involvement in multiple diseases [58–60]. Herein, we propose that the CLMS techniques, though, which have not been previously carried out on such problems, could be merged with novel interactomics techniques (proximity-dependent labeling by BioID or an engineered biotin ligase by TurboID technique). Combining these two approaches may bring spatial resolution to CLMS at a sub organelle level, since the BioID radius is estimated ~10 nm and CLMS is aimed on the proteins in the neighborhood of a given bait.

While there has been some success in the proteome wide CLMS [61–64], in the case of both chemical cross-linking MS and the related techniques of native MS [65,66], a prior purification of single protein or protein complex is typically necessary, and this can be acquired through over expression with the purification of a recombinant version of particular system. When investigating the protein assemblies, individual components need to be purified to reconstitute the whole complex later *in vitro* and, alternatively, *in vivo* reconstitution can occur by co-expressing various subunits [67]. Because the reconstituted systems mostly benefit from large yields to aid the structural analysis and they are frequently done using bacteria (mostly *E. coli*) as the host system, functionally important post-translational modifications and interacting protein partner associations may be lost during this process. Thus, various biochemical approaches must be explored when beginning a new project in order to directly enable the isolation of endogenous protein complexes from cells or tissues [68].

Modern mass spectrometry that is coupled with the chemical cross-linking of juxtaposed amino acids can provide important structural information. Two main types of the cross-linking strategies involving either the activation of the cross-linking reagent by UV or chemical methods to enable cross-linking [11–13,69]. Chemical cross-linking is a classical approach for determining protein–protein interactions and is also one of the first approaches that has been used to map large complexes, for example, the ribosome [2]. Generally, the cross-linking techniques link two or more proteins present in a complex by covalent bonds and, as the name implies, via a molecule designed to bridge juxtaposed amino acids, i.e., to chemically cross-link residues. The chosen cross-linker is a chemical reagent that contains two or more reactive groups connected through a spacer or linker of various lengths [70]. By using this method, low-affinity protein–protein contacts, or some specific interactions, can be detected that are difficult to characterize by other methods (e.g., nuclear magnetic resonance (NMR), X-Ray, etc.). Moreover, the cross-linking techniques have also been applied to stabilize transient protein–protein interactions in a dynamic process both *in vitro* and *in vivo* [1,2]. However, there can be considerable weakness in these chemical methods that are related to the lack of spatial localization in a cell and a lack of control over activity. Thus, to evaluate PPIs as similarly as possible to the native conditions, photo-cross linking methods are valuable due to their ability to generate reactive species *in situ* instantaneously by irradiation with UV-light [11–13,69].

Cross-linking is always followed by other downstream methods to further analyze the cross-linked proteins, often using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to separate the cross-linked from non-cross-linked proteins, tandem affinity

purification (TAP) or immunoaffinity chromatography for affinity-based purification of cross-linked products, and mass spectrometry methods for the interacting partner identification [71,72]. Additionally, SDS-PAGE analysis is very helpful in the early stages of analysis when empirically working out the correct ratio of cross-linker to protein complex. A limitation of using chemical cross-linking while using these methods is the high risk of detecting non-specific interactions. However, these limitations can be addressed using more than one cross-linker of differing activities or spanning various distances and by varying the ratio of reagent to protein complex. Non-specific interactions can result from proteins in close proximity that may not be functionally related. These suggest that while the CLMS technique is relatively straightforward to implement, the identification of relevant cross-linked proteins could be quite demanding due to the intracellular dynamic range of expression of proteins, which can range from one to one million copies low abundance of cross-linked species [1].

Innovative developments in the biological applications of MS led to the development of a large number of methods, and it has made it relatively simple to identify proteins alone or in complexes using CLMS technique. Large macromolecular complexes like ribosomes or exosomes, have been purified and analyzed directly using mass spectrometry [73–76]. Recently, chemical cross-linking combined with mass spectrometry based structural techniques have hit their stride allowing for various biologically relevant molecular machines to be successfully studied in the past few years using this combination [77]. The various workflows that have been developed to implement CLMS represent a vast toolkit that can help to provide novel insight into the structure and organization of proteins in order to define protein–protein interactions and probing PPI interfaces.

2. Concept and Perspectives of Cross-Linking Mass Spectrometry

In the CLMS approach, chemical cross-linking reagents are used to join the components of interacting complexes, followed by LC-MS/MS (Liquid chromatography-tandem mass spectrometry) analysis that enables *in vivo* and *in vitro* methods to define the native PPIs of a protein complex, under optimal conditions. The visualization of the interacting regions allows distance maps within the protein complexes or within the protein to be created, e.g., low resolution three-dimensional maps of the interactions can be generated. In addition, to be valuable for defining protein–protein interactions, CLMS has emerged as a technique for interactomics and the structural biology of multi-protein complexes. For example, the CLMS based methods are now capable of capturing protein–protein interactions from their native environment, uncovering physical interaction contacts between them and, thus, providing the determination of both identity as well as connectivity of PPIs in cells (Figure 2) [78,79]. Some of the advantages of using the CLMS technique over traditional structural methods, like X-Ray diffraction, cryo-EM, or NMR include: limited amounts of starting material are required (i.e., typically only nanograms), no need for exceptionally pure protein(s), and much more rapid turnaround in the workflow allowing for more rapid hypothesis generation and testing. Those traditional approaches (*vide supra*) are limited by the particular proteins that can be easily expressed or crystalized, whereas the CLMS has an ability to examine the biologically relevant interactions that are close to the physiological state of an organism [77–79]. Moreover, PPIs must be very strong to survive the condition of extraction and purification required by the general sample preparation in traditional methods, and this makes it very challenging to analyze interactions in their native state.

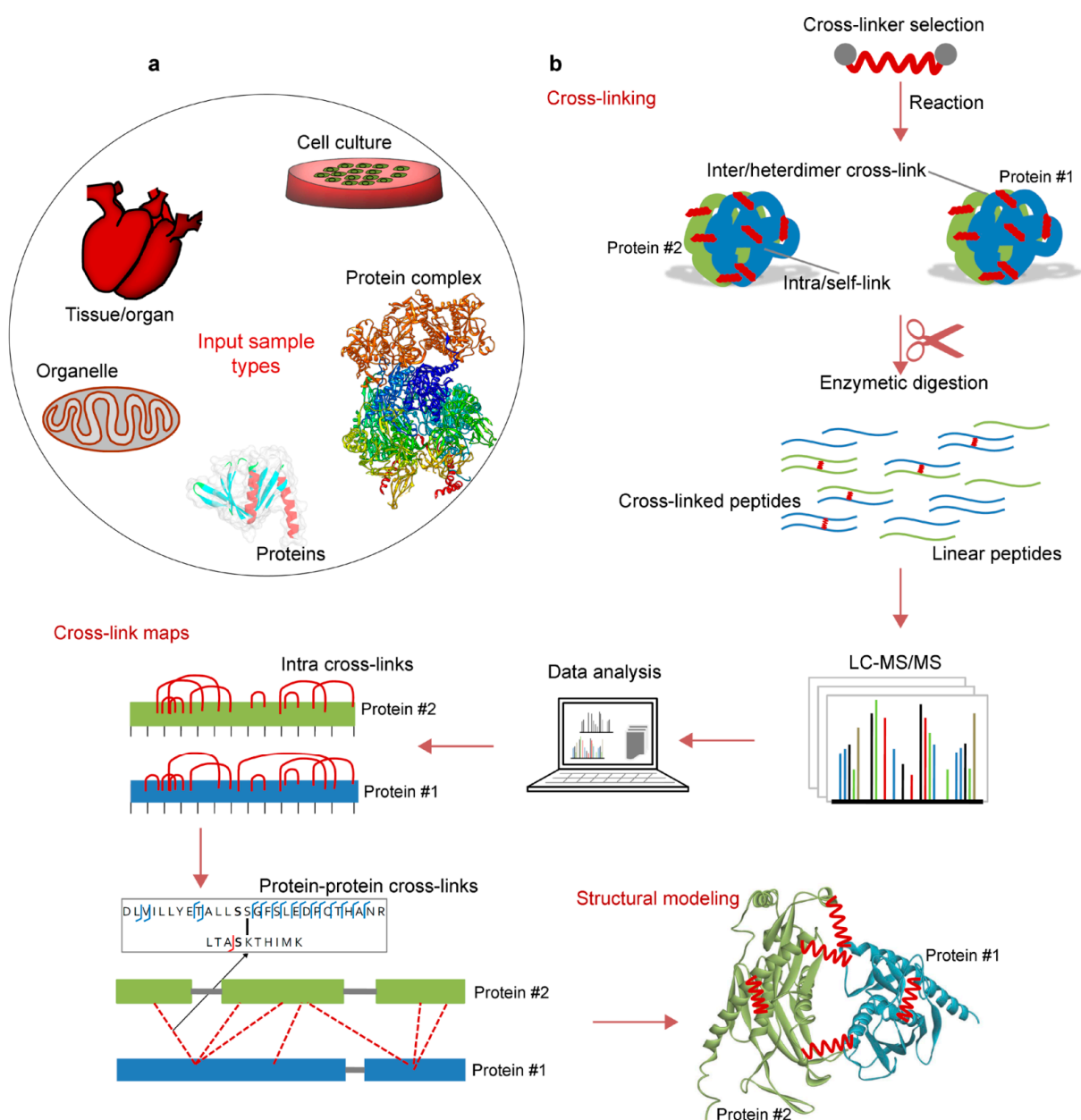


Figure 2. The generic workflow of cross-linking mass spectrometry (CLMS) experiments. (a) A wide range of samples (from purified proteins to intact tissues and organs) applicable to perform CLMS experiments [11,80]. (b) In a representative CLMS workflow, a selected linker is applied to the sample and the cross-linking reaction is carried out. Depending on the actual chemistry of interest, the reaction is stopped through chemical quenching or removal of the reagents. The proteins can be then digested in solution or gel to produce a mix of cross-linked and linear peptides. Prior to mass spectrometry analysis, the cross-linked peptides are often enriched by chromatographic methods, for example, the size exclusion chromatography, ion exchange chromatography, or purification through an affinity tag. Finally, the sample is subjected to LC-MS/MS (Liquid chromatography-tandem mass spectrometry) acquisition pipelines that have been developed to add the likelihood of selecting cross-linked peptide precursors for fragmentation. Using a variety of search software, two linked peptides can be identified from spectra and through the methods determining the false discovery rate, the list of matches can be filtered to the desired confidence. The cross-links can be also visualized by integrative modeling techniques. CLMS techniques convey the structural information in the form of distance restraints on single protein, protein complexes and allows to portray protein networks [11,80–82]. In this figure the structures of protein or peptides were prepared using BIOVIA Discovery Studio (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) visualizer tool.

2.1. Associate Methods for CLMS Technique

Three main chemical approaches that involve MS analysis on the peptide level are: (i) the exchange of labile hydrogen atoms with deuterium atoms in hydrogen/deuterium exchange (H/D exchange or HDX) methods, (ii) the covalent modification (i.e., painting or surface modifications) of amino acid residues (mostly the functional groups in side chains) in various covalent labeling workflows, and (iii) chemical cross-linking, in which two spatially proximate amino acid side chains are covalently coupled (Figure 3) [72]. Particularly, the CLMS methods are complementary to other MS techniques for structure analysis. As an example, the HDX-MS can provide information regarding the regions that may be analyzed by CLMS and it is frequently used to examine the conformational flexibility of protein complexes. Following this, CLMS can be used to obtain distances between interacting regions and fill in information that might be missed by traditional method, e.g., regions of NMR or X-Ray data that are poorly defined (Figure 3). When considering the fact that the CLMS workflows often involve simply injection peptide mixtures, as would be done for any shotgun proteomic experiment, there need not be any significant interruption to overall laboratory workflows apart from upstream use of specific cross-linking.

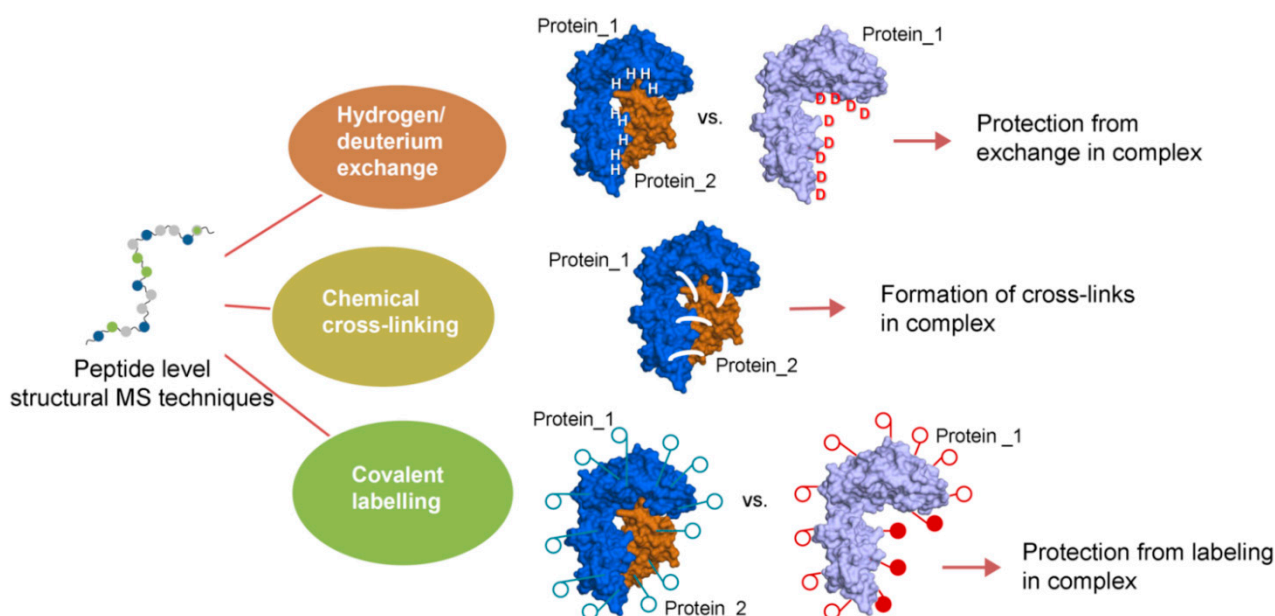


Figure 3. Three important mass spectrometry approaches used in the structural proteomics analysis of binary complexes. In the hydrogen/deuterium exchange, H_2O is replaced by D_2O and the resulting exchange associated with a mass increase can be detected by mass spectrometry (MS). Chemical cross-linking, comprises the covalent coupling of two reactive groups within a protein or between two different proteins, and then by introducing cross-links at the specific residues, spatial information at different levels can be obtained using MS. In covalent labeling, irreversible modifications are introduced at the reactive side chains and the solvent or surface exposed residues in proteins can be identified [23,72]. In this figure the protein or peptide structures were prepared using the BIOVIA Discovery Studio (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) visualizer.

Despite the growing popularity of CLMS technique, there remain limitations that need to be overcome in order to make it more successful in defining the *in vivo* state of PPIs. Its main limitation is the small depth of interactome coverage [61]. To date, the 'maximum number of cross-links identified in system-wide studies is ~10,000, however this number is expected to grow as methods continue to improve. Specific challenges to overcome full interactome coverage are the high dynamic range of expression of proteins (from 1 to 10^6 intracellularly) and diversity in their binding affinities. Additionally, these include the variety of amino acid residues that can be targeted by cross-linkers and the

decreased solubility of cross-linked protein complexes that, in turn, causes problems during digestion and the ionization of large cross-linked products [11]. Addressable challenges include: (i) reducing sample complexity by liquid- and gas- phase methods, which, in turn, can increase detectable dynamic range, and thus, detect lower copy number proteins, (ii) targeting additional functional groups in proteins in an aqueous environment, and (iii) most CLMS methods target amine reactive cross-linkers that leads to a 'dark' interactome that is blind to PPIs, where the primary amine groups are absent or scarce in the contact region [11,83,84].

Identifying the potential regions of protein intra- and inter- molecular interactions is considered the main power of CLMS in part, because it can also provide direct information on a range of distance constraints that can be used to develop or enhance three-dimensional models of protein in complex structures [85,86]. However, rather than CLMS data being an end unto itself it is used as one of a set of analytical methods that, together, provide structural insights. In the case of CLMS, the main challenges are to achieve robust workflows that enable a comprehensive capture of dynamic biological systems interactions in their native environments in a routine manner [11]. Typically, the first step in any CLMS experiment is to define the chemical cross-linking reagents, which is based on known or expected amino acid sequence, be used, and then next to define the ratio of reagents to proteins. Inducing too many cross-links may make the complex undigestible, while too few may not provide useful information. This ratio is usually empirically defined by the following changes to the complex by SDS-PAGE. After this ratio is developed, the cross-linked proteins are digested to peptides using an appropriate enzyme. In some cases, an enrichment step may be integrated to isolate cross-linked peptides from the overwhelming number of non-cross-linked peptides in the sample that prohibits the detection of the lowest copy number cross-linked peptides. Finally, the peptides are introduced into an LC-MS/MS instrument for data collection, and specialized software packages designed for CLMS studies are then used for data interpretation (Figure 2) [72].

2.2. Chemical Cross-Linkers Structure and Chemistry

A number of options for chemical cross-linking reagents with a wide variety of reactivities are available; they contain reactive ends to a variety of chemical specificity and these reagents are used to cross-link two regions within a protein (intramolecular) or between two different proteins (intermolecular cross-link; homo- or hetero- dimer). These intramolecular cross-links can stabilize tertiary or quaternary structures of a protein, and intermolecular cross-links stabilize the protein–protein complex (the main interest in the field of PPIs analysis) [87]. The growth in the use of protein CLMS has led to a wide variety of cross-linking reagents, with various distance constraints between functionally reactive groups. It may be required that more than one reagent each with different distance constraints could be used to aid construction of the three-dimensional protein structure or PPIs (Figure 4) [87,88].

In the simplest terms, chemical cross-linking reagents are comprised of two reactive functional groups that are separated by a spacer arm, which defines the distance between functional groups (Figure 4a). For an intramolecular (within a protein) cross-linking, a short spacer arm is more likely to be useful, whereas, for the intermolecular cross-linking studies, longer spacer arms can be more efficient. In addition, this difference is necessary to satisfy the steric constrain effects that guide the distance between potential reaction sites for cross-linking. The cross-linkers can be homobifunctional (having the same reactive functional groups at both ends of the spacer arm) or heterobifunctional (with different reactive groups at either end of the spacer arm). Homobifunctional cross-linkers are ideal for capturing a protein–protein interaction snapshot, while the use of heterobifunctional cross-linkers allow for two-step sequential conjugations that minimize undesirable polymerization or self-conjugation [87–89]. Overall, the chemical reactivity is the basis for choice of a given cross-linker. Different chemical properties that facilitate their use for specific applications and affect choice can be made by meeting specific criteria,

including target functional groups, cell membrane permeability, and solubility. Cross-linking reagents are classified based on some general features, like chemical specificity (homo- or hetero- bifunctional structure), spacer arm length, water solubility, and cell membrane permeability (i.e., whether it is desired for the reagent to permeate cells or cross-link hydrophobic proteins within membranes), and instantly reactive or photoreactive groups [87–89].

The ideal cross-linkers for MS analysis with a focus on the different aspects of chemical structure and reactivity contain molecules with functional groups, by which at least two of the groups are reactive and capable of conjugation (Figure 4a). These types of cross-linkers are clearly the homobifunctional amine (lysine)-reactive N-hydroxysuccinimide (NHS) and sulfo-NHS esters (e.g., disuccinimidyl suberate, DSS; and, bis(sulfosuccinimidyl)suberate, BS3; Figure 4b and Table 1) [71,90]. Targeting the lysine residues is desirable due to their relatively high prevalence (~6% of all residues), their distribution across solvent-accessible protein surfaces, and the specificity of primary amine-targeting chemistries. Because it can be difficult to characterize the proteins with few or no lysine residues using these amine-reactive agents, other residues, like serine, threonines, and tyrosines, are also targeted for cross-linking [71,80]. The majority of recent studies have made use of non-cleavable cross-linkers and, thus, the DSS and BS3 have become the reagents of choice. The only difference between these cross-linkers is a sulfonic acid group that is incorporated into BS3 to improve water solubility, and to bridge a distance of 11.4 Å, which results in Cα–Cα distance of ~27 Å [71]. The tremendous success of DSS or BS3 reagents is due to their simplicity, reaction specificity, ease of use, reaction product stability, and lack of reaction by products. Their successful use has been demonstrated by integrative structural biology in purified complexes, in organelles, in cells, and labeled, as well as label-free quantitation for analyzing structural changes of protein [89]. Expanding further, MS-cleavable cross-linkers, such as disuccinimidyl sulfoxide (DSSO) and disuccinimidyl dibutyric urea (DSBU), are also widely being used, since they provide an additional level of information in their tandem mass spectra that contain characteristic fragment ions generated during tandem MS experiments. The principal spacer lengths of cross-linkers in the community-wide CLMS study is in range of 10–12.5 Å, based on the most preferentially used cross-linkers, cleavable BS3 and DSS, as well as non-cleavable DSSO and DSBU (Figure 4 and Table 1) [71,91].

Table 1. Few examples of a diverse set of chemical cross-linkers supported by different packages or tools [105–112].

Tools	Chemical Cross-Linkers Supported	Website
MeroX (StavroX included) [106]	BS2G, BS3/DSS, BS3/DSS-D0/D12, CDI, DC4, DSAU, DSBU, DSSO, DST, EDC, Formaldehyde(12), Formaldehyde(24), SDA	http://www.stavrox.com/Download_MeroX_Win.htm (accessed on 23 February 2021)
Spectrum Identification Machine for Cross-Linked Peptides (SIM-XL) [107]	DSS, DSG, DSSeb, DSS/DSG/DSSeb (with reporter ions only), XPLex C6N2, XPLex C3N2, XPLex C6Ac2, XPLex C3Ac2, Disulphide, zero-length	http://patternlabforproteomics.org/sim-xl/ (accessed on 23 February 2021)
Xilmass [108]	DSS (d0/d12), BS3(d0/d4), EDC and GA	http://compomics.github.io/projects/xilmass.html (accessed on 23 February 2021)
xQuest/xProphet [109]	BS3, DSS etc.	http://proteomics.ethz.ch/cgi-bin/xquest2.cgi/download.cgi (accessed on 23 February 2021)
XiSEARCH [110]	BS2G, SDA, BS3, DSSO, EDC, NonCovalent, Linear Search	https://www.rappsilberlab.org/software/xisearch/ (accessed on 23 February 2021)
Kojak [111]	BS3, DSS etc.	http://www.kojak-ms.org/param/cross_link.html (accessed on 23 February 2021)
pLink 2 [112]	BS2G, BS2G_heavy, BS3_heavy, DSS, EDC-DE	http://pfind.ict.ac.cn/software/pLink/ (accessed on 23 February 2021)

In addition to tools presented in this table other programs such as; XLSearch, ECL 2.0, MaxQuant, xTract, Protein Prospector, pQuant, mMass, CLMSVault, xVis, XlinkX, Mango, CLPM, Crux, DXMSMS, FINDX etc., can also be used for identification of cross-linked peptides [11,80,105].

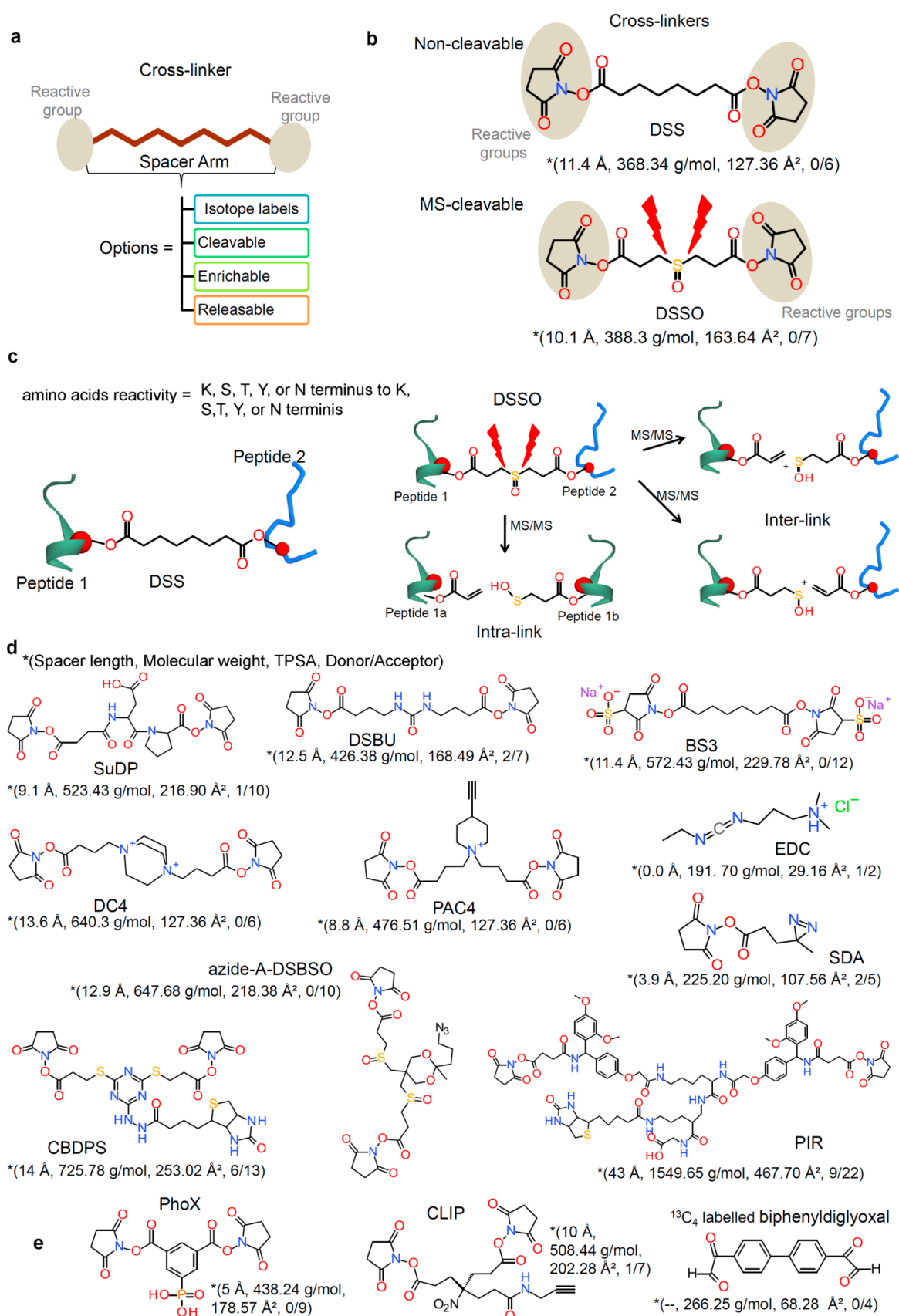


Figure 4. Cross-linking reagents anatomy, chemistry, and evolution of cross-linkers. (a) Cross-linking reagents comprise various reactive groups, spacer scaffolds with different lengths. Based on the applied experimental workflow, a variety of

isotope labeled, cleavable, enrichable, releasable, and spacer groups are in use. A large number of cross-linkers establish their combinatorial complexity and further supports a larger number of future cross-linkers [87–89]. (b,c) Structures of the commonly used cross-linkers for analyzing protein–protein interactions involving non-cleavable (disuccinimidyl suberate, DSS) and MS-cleavable (disuccinimidyl sulfoxide, DSSO) cross-linkers [71,91]. Representation of the suggested fragmentation schemes (interlink as well as intralink cross-linked peptides) of DSS and DSSO cross-linked peptides. (d) In addition to DSS and DSSO, most commonly used cross-linkers are shown (SuDP, Disuccinimidylsuccinamyl aspartyl proline; DSBU, Disuccinimidyl dibutyric urea; BS3, Bis(sulfosuccinimidyl) suberate; DC4, 1,4-Bis[4-[(2,5-dioxo-1-pyrrolidinyl)oxy]-4-oxobutyl]-1,4-diazoniabicyclo[2.2.2]octane; PAC4, 1,1-Bis[4-[(2,5-dioxopyrrolidin-1-yl)oxy]-4-oxobutyl]-4-ethynylpiperidin-1-ium; EDC, 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide; Azide-A-DSBSO, Azide-tagged, acid-cleavable disuccinimidyl bissulfoxide; SDA, succinimidyl 4,4'-azipentanoate; CBDPS, Cyanurbiotindipropionyl succinimide; and PIR, protein interaction reporter) [71,72,89,91–100]. (e) For the enrichment of cross-linked products/peptides a few specific cross-linkers are used such as: PhoX linker (non-cleavable but containing a phosphonic group) [101], CLIP cross-linker (introduces alkyne groups to cross-linked peptides) [102], and ^{13}C labeled form of biphenyldiglyoxal (diglyoxal cross-linkers) [103,104]. For the cross-linkers on panel b, d, and e, different properties such as spacer length, molecular weight, topological polar surface area (TPSA), and number of hydrogen bond donor or acceptor atoms in the cross linker are described. TPSA and hydrogen bonds were computed using the Molecular Operating Environment (MOE; Chemical Computing Group Inc., Montreal, QC, Canada) package, and the protein/peptide/chemical structures were prepared using BIOVIA Discovery Studio visualizer and BIOVIA draw tools (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA).

The spacer arm of cross-linkers can also act as a scaffold for functionalities that focus on the low abundance and complexity of analytes, as well as characteristics of cross-linking products that require analysis. Such functionalities involve MS-cleavable groups, isotope-coding, enrichment handles, and related capture, as well as release groups. Modular design approaches have been applied to cross-linker design and synthesis to generate multiple functional groups. For example, this concept is embodied by the popular protein interaction reporter (PIR) cross-linkers [89,92]. A class of PIR cross-linkers, containing MS-cleavable reagents that can also be enriched through a biotin label [93]. In 2006, a MS-cleavable cross-linker SuDP (Disuccinimidylsuccinamyl aspartyl proline) containing a labile aspartate-proline bond was presented [94]. In principle, the cross-link identification strategy that was presented for SuDP [95] formed a basis for the consecutive development of proteome-wide CL-MS workflows by using other MS-cleavable linkers. Various other MS-cleavable linkers, such as the DSBU [96], DSSO [83], CBDPS (Cyanurbiotindipropionyl succinimide) [97], and DC4 (1,4-Bis[4-[(2,5-dioxo-1-pyrrolidinyl)oxy]-4-oxobutyl]-1,4-diazoniabicyclo[2.2.2]octane) [98], are revealed, as well as applied in different protein systems (Figure 4). As derivatives of DSSO, two different trifunctional cross-linkers were designed harboring azide (azide-A-DSBSO: azide-tagged, acid-cleavable disuccinimidyl bissulfoxide) or alkyne (alkyne-A-DSBSO) groups, which enable affinity purification approaches that are based on click-chemistry [99]. Likewise, the trifunctional PAC4 (1,1-Bis[4-[(2,5-dioxopyrrolidin-1-yl)oxy]-4-oxobutyl]-4-ethynylpiperidin-1-ium) linker was designed from DC4 having an alkyne group for an affinity enrichment of cross-links (Figure 4) [71,100]. The impressive array of chemistry applied for the enrichment of cross-linked products is provided by the PhoX linker (Figure 4e) [101]. The PhoX linker is non-cleavable, but contains a phosphonic group, which allows for the cross-linked products to be enriched through routine methods, like titanium dioxide substrates, which are used for phosphopeptide enrichment by immobilized metal ion affinity chromatography (IMAC) [71]. The purified cross-linker with the PyrR target protein complex was incubated to evaluate the diglyoxal-based (^{13}C labeled biphenyldiglyoxal) cross-linkers (Figure 4e) and confirm their use to identify protein–protein interactions [103,104]. The work by AN Holding [103] described that these reagents (biphenyldiglyoxal; Figure 4e) were useful for recognizing the formation of arginine-selective cross-links in the human PyrR complex.

2.3. Cross-Linked Sample Preparation for MS (Protein–Protein) Analysis

Protein cross-linking produces a challenging task for a mass spectrometrists, due to an introduction of multiple interconnected amino acid sequences into a mass spectrometer. Cross-linked proteins could be analyzed by either top-down or bottom-up mass spec-

trometry. Less frequent top-down mass spectrometry measures intact protein complexes; however, more common bottom-up mass spectrometry relies on simpler peptides that need to be generated prior MS measurement via proteolytic cleavage. Herein, we shed light on possible sample preparation pipelines used to turn a cross-linked protein sample into a clean peptide solution that is introducible into MS, and discussed the major factors impacting the detection of cross-linked peptides.

The peptide sample preparation process has a major impact on subsequent cross-linking experiments using mass spectrometry, because the investigation of protein interacting domains directly relies on peptides detectable by MS. The detection of peptides by MS is influenced by the chemicals used in sample preparation that may interfere with the eventual ability of any given peptide to be detected by ionization, and directly by the peptide's chemical properties that make it more or less ionizable. The very first step after cross-linking is to ensure the cross-linked protein remains in solution, which can be a problem for some proteins that may have been marginally soluble prior to cross-linking. This decrease in the solubility may require the addition of chemicals, like detergents, to the buffer to improve solubility or simple dilution with lysis buffer. For all of these reasons, it is important to use mass spectrometry compatible buffers if downstream analysis does not or cannot include a buffer changing step. Rapid and effective cross-linked peptide detection also depends on the proteolytic digestion protocol. There are several cross-linked protein digestion protocols that have been reported (Figure 5).

Commonly, cross-linked proteins are electrophoretically pre-purified from organic contaminants and aggregates while using SDS gel electrophoresis, followed by the extrusion of peptides from the gel band after proteolysis in the same gel band. A study has been reported that focused on QCLMS [113], in order to investigate the protein structural conformations in solution used with in-gel digestion protocols (Figure 5). The digestion of photoactivated cross-linked proteins was performed after SDS pre-separation and concentration in polyacrylamide gel. Müller et al. claimed that they were able to detect 414 unique residue pairs, out of which 292 (70%) were quantifiable across triplicate analyses with a coefficient of variation (CV) of 10% [113]. In addition, several other studies suggested the use of 'in-gel' digestion in their cross-linking proteomics experiments (Figure 5) [114–118]. Nevertheless, the 'in-gel' digestion has the potential for the loss of cross-linked peptides, where their yield relative peptides without cross-linking is low. The main factors contributing to decreased cross-linked peptide recovery from a gel include poor peptide solubility, bulkiness due to the cross-link, and entrapment in gel pores due to peptide branched structure (Figure 5). Therefore, Petrotchenko et al. [119] created an 'out-gel' tryptic digestion procedure for chemical cross-linking studies with mass spectrometric detection. This out-gel digestion procedure is based on SDS-PAGE separation, followed by the passive diffusion of cross-linked proteins from the gel. Cross-linked protein digestion takes place outside the gel, which increases the probability that cross-linked peptides will be detectable (Figure 5). Petrotchenko et al. included strong cation exchange (SCX) chromatography, followed by the zip-tip cleanup method on the C18 reversed-phase media to remove contaminants and salts from sample prior MS. Moreover, they showed that 93% of the cross-links have better or equal recovery while using 'out-gel' tryptic digestion, as compared to the standard 'in-gel' tryptic digestion [119].

The in-solution approach is another alternative that leads to improved cross-linked peptide recovery as compared to 'in-gel' digestion. It is important to ensure that the sample does not contain mass spectrometry incompatible detergents, and unwanted protein aggregates prior to 'in-solution' digestion (Figure 5). Mass spectrometry incompatible detergents and other organic contaminants will remain in the peptide solution, and they might interfere with peptide separation or cause ion suppression. Parfentev et al., in their study of the 'n²' problem of cross-linked peptide search, used 'in-solution' digestion to prepare a model of cross-linked proteomic data [120]. The 'n²' problem represents a challenge that consists in selecting any residue that is capable of cross linking, as a candidate to be cross-linked to any peptide considered in a specific experiment. Therefore, the 'n²' problem creates

$(n^2 + n)/2$ possible cross-links for 'n' number of peptides [121]. There have been many other reports relying on 'in-solution' digestion of cross-linked peptides [122–126]. A vital alternative to 'in-gel' and 'in-solution' digestion protocols, especially when the cross-linked sample contains detergents and contaminants, is 'Filter Aided Sample Preparation' (also termed FASP) method. The FASP method (Figure 5) employs a mass cut-off molecular filter allowing for the high recovery of cross-linked peptides after proteolytic digestion parallel with removal of organic contaminants. Rey et al. reported the use of eFASP protocol to digest membrane proteins cross-linked using trifunctional cross-linker, named 'NNP9', in the presence of a MS incompatible detergent [127]. In addition, they suggest that using eFASP followed by the enrichment of cross-linked peptides on monoavidin beads, leads to a drastic improvement in the number of identified cross-linked peptides, when compared to standard gel based digestion [127].

The protease or multiple proteases used to digest the cross-linked protein must be carefully chosen along with a digestion protocol, as this presents yet another factor that could substantially enhance or diminish the result of the experiment. This is in part due to limitations in mass spectrometry, and the typical use of acidic solutions to detect basic peptides by MS. Specifically, peptide length and physicochemical characteristics influence MS detectability as does the co-eluting non-peptide matrix of chemicals used for sample preparation. Recently, it has been shown that the detectability of large tryptic peptides could be enhanced by including stepwise multi protease digestion. Mendes et al. reasoned that sequential digestion could offer an access to the sequence space that otherwise would remain unseen, therefore they employed additional proteases such as AspN, LysC, and chymotrypsin. Multiple protease digestion could reveal mechanistically important protein regions that would not be detectable with tryptic digestion alone [110].

Cross-linked peptides represent only a relatively small portion of the peptide pool in any given cross-linked protein sample after proteolysis. Therefore, sometimes cross-linked peptide enrichment or peptide pre-fractionation steps are included. Ion exchange chromatography represents a potent tool to separate linear from cross-linked peptides, as cross-linked peptides have a higher isoelectric point and are naturally more likely to have higher charge states. Fritzsche et al. demonstrate the benefit of pre-separation after cross-linking by comparing the number of identified cross-linked peptides with and without prefractionation step. Their data clearly show the benefit of introducing SCX, which is reflected in a rapidly increased identification of interpeptidal cross-linking products and overall gain in structural information [128]. The ChaFRADIC (charge-based fractional diagonal chromatography) protocol has been developed to enhance SCX separation of linear from cross-linked peptides. In principle, it relies on blocking free primary amines with dimethyl chemistry prior to proteolysis. Following this reaction, cross-linked proteins are digested with trypsin and resulting peptides are separated by SCX into fractions with theoretical charge of +1, +2, +3, or +4. Subsequently, the peptides are trideutero-acetylated to block newly created N-termini. Then SCX is performed again and the net charge of internal peptides will be reduced by one, while net charge of N-terminal peptides will not be changed as they have already been dimethylated. Internal peptides shift to earlier fractions while N-terminal peptides elute the same as before [129]. Introducing SCX pre-separation step is popular in cross-linked proteomics as it is well documented in several studies (Figure 5) [119,128–130]. Nevertheless, SCX fractions require additional desalting steps prior mass spectrometry analysis which might lead to extensive peptide losses.

Alternatively, the cross-linked proteins might be enriched by pull-down via a biotin group or various other groups introduced during cross-linking reaction. Chowdhury et al. developed a novel multifunctional CLIP cross-linker (Figure 4e) that introduces alkyne groups to cross-linked peptides. This group could be exploited to enrich cross-linked peptides via alkyno-azide chemistry after the reaction [102]. Tan et al. developed trifunctional linker cross-linking free amino groups of interacting proteins. Their cross-linker contained a biotin tag for cross-linked peptide purification. Interestingly, the cross-linker also possessed a chemical cleavage site to detach biotin tags after purification. As needed a spacer arm can

be included to introduce isotope-labels for quantitative purposes [131]. However, affinity purification of cross-linked products enriches the sample for peptides that reacted with cross-linking reagent, but it does not distinguish between linear and cross-linked peptides. Therefore, exploiting the size difference between linear and branched cross-linked peptides is another option of how to deal with low abundance of cross-linked peptides. Several studies employed a SEC (size exclusion chromatography) step to pre-separate emerging peptides based on their size [132,133]. Leitner et al. assembled a cross-linking workflow including SEC chromatography separating true cross-linked proteins from all kinds of linear proteins that does not contribute to a description of protein complex. Moreover, they demonstrate that the identification of cross-linked proteins increases 3-fold upon introduction of SEC [133]. Adopted workflow was then applied to enhance study of cross-linked protein complexes of human protein phosphatase 2A (PP2A) [134].

Identification of protein cross-links depends on additional factors such as the cross-linked peptide ionization efficiency, mobile phase composition, electrospray setting, mass analyzer, acquisition method, and data analysis. The cross-linked peptides could be eluted with other peptides (non cross-linked) or even with the contaminants, and therefore they will compete for a given number of protons or charges in the electrospray source. Peptides with better ionizability will tend to charge more and peptides with low ionization efficiency will be suppressed resulting in no or low signal. Moreover, gold standard DDA (data dependent acquisition) methods can detect only a limited set of charged precursor peptides at a given time. The DDA methods pick ions for tandem MS in stepwise fashion from the most abundant ion at any given time, which naturally excludes lower abundance ions. It is not uncommon that more peptides elute at the same time than can be selected for tandem MS, which could lead to complete exclusion of low intensity precursors ions, some of them may be cross-linked peptides [135–138]. Successful cross-linked peptide detection is only the very first step of a cross-linked experiment and always includes finely tuned downstream data analysis, which is another important step in the overall pipeline to identify juxtaposed amino acids.

2.4. Cross-Linking by UV for Protein–DNA Interactions

Protein–DNA interactions, fundamental to the functionality and stability of the genome, control essential cellular processes like replication, transcription, repair, and recombination. To understand the DNA-dependent processes, mapping of such protein–DNA interactions as well as identification of specific sites of interaction are required [139]. Many of the previously described CLMS workflows can be adopted to obtain such insights as has been recently demonstrated by Stützer et al. [140], who performed mass spectrometric identification of proteins interacting directly with DNA in reconstituted and native chromatin after cross-linking by UV light. Analysis of contact interface at amino acid level was possible by this approach, and they also described the possible means to distinguish the protein–DNA and protein–RNA interactions by performing a single experiment [140]. Chromatin is one of the most prominent protein–DNA complexes of a eukaryotic cell and in this, a core of eight histone proteins (2xH2A, 2xH2B, 2xH3, and 2xH4) associate with DNA in a repetitive manner to facilitate structural and functional organization of the genome [139]. Several studies described that RNA binding sites in proteins can be detected efficiently by using UV cross-linking combined with mass spectrometry [141,142]. Similarly for the cross-linking of protein and DNA components, a well established protein–DNA cross-links are induced in vivo after the exposure of cells to UV light, ionizing radiation or alkylating agents, in order that lead to bulky DNA lesions [140,143]. UV irradiation of DNA triggers a cellular cascade called DNA damage response, comprising a multitude of proteins and likewise, this UV irradiation has been successfully applied to cross-link single- and double- stranded DNA to proteins for analyzing chromatin dynamics [144]. Accordingly, the UV cross-linking with mass spectrometry could be beneficial to explore structural and functional relations in protein–DNA systems [140].

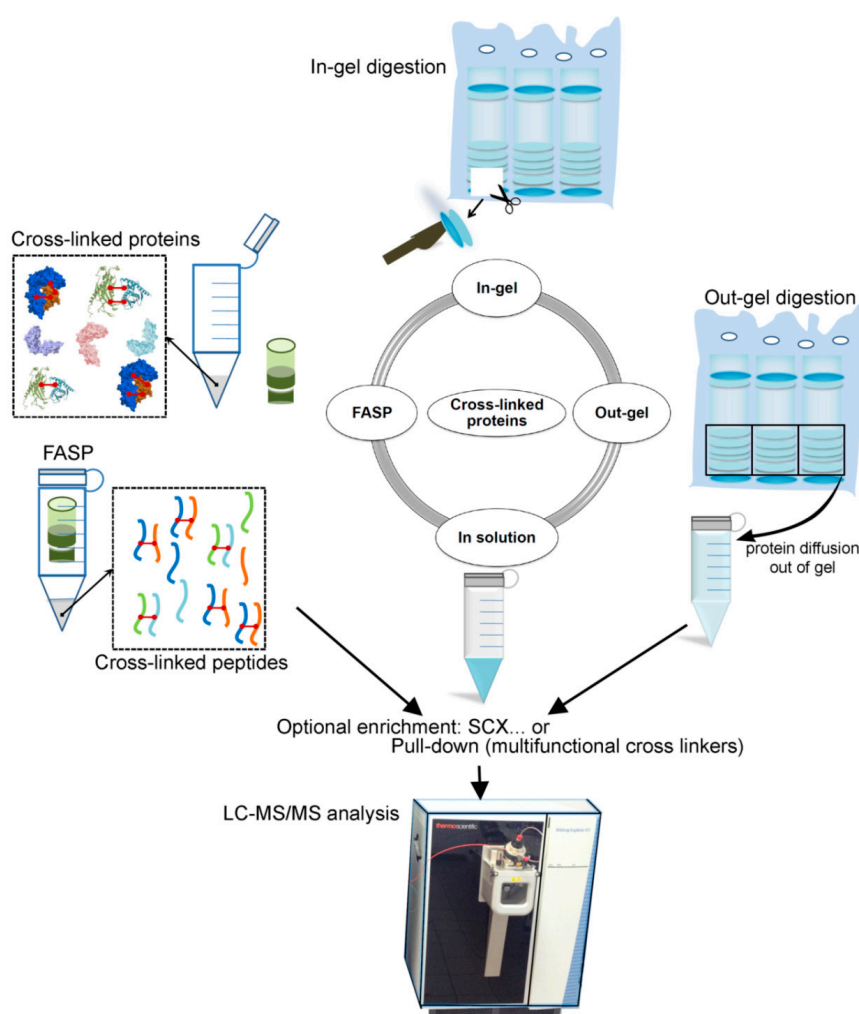


Figure 5. A workflow of determining protein structure, as well as the protein–protein interactions using the most frequent bottom-up cross-linking mass spectrometry. Extracted cross-linked proteins could be turned into peptides by several digestion protocols and proteases. Protein digestion protocol is chosen based upon up-stream steps taken to prepare cross-linked protein extract. FASP (filter aided sample preparation), ‘in-gel’, and ‘out-gel’ digestion protocols are used predominantly when a sample contains mass spectrometry incompatible substances, e.g., detergents or contaminants. However, ‘in-solution’ digestion might be used along with proteomic samples, which are already mass spectrometry compatible and contain predominantly cross-linked proteins of interest. Complexity of the cross-linked peptide samples might be reduced by separating cross-linked peptides from linear peptides after digestion. Difference in physico-chemical properties of cross-linked and linear peptides is often exploited to perform strong cation exchange or separation based on difference of peptide size. Alternatively, cross-linked peptides could be pulled down if a linker containing an affinity tag such as for example, biotin group has been used. Down-stream mass spectrometry analysis is frequently done in data-dependent acquisition mode (DDA). In the figure, representation of protein or peptide structures were prepared using BIOVIA Discovery Studio (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) visualizer.

Recently, Stützer et al. adapted the established protein–RNA CLMS workflow for detecting protein–DNA cross-links [140]. In their initial experiments they showed that UV irradiation at 254 nm efficiently cross-link histone proteins to double stranded DNA. Consequently, they set out a protein–DNA cross-linking workflow that can be used for simple protein–DNA complexes, for example; oligonucleosomes, chromatin binding factors, and also complex systems like cell nuclei. In this workflow [140], first the linker histone–DNA complexes along with single nucleosome and 12 mer oligonucleosomal (chromatin) arrays containing *Xenopus laevis* core histones [145] were prepared. After irradiation, protein and DNA complexes were hydrolyzed with DNA nuclease and trypsin to generate peptides,

oligonucleotides and cross-linked peptide-DNA oligonucleotide conjugates acceptable to MS analysis. Oligonucleotides without cross-links were excluded by C18 reversed phase chromatography and finally, peptide-DNA conjugates were enriched by using TiO₂ affinity chromatography (Figure 6a) [141,146]. Purified peptide-DNA oligonucleotide conjugates were then analyzed through LC-MS/MS, and resulted MS data were analyzed using the RNP^{xl} computational workflow [141,147] in the OpenMS (<https://www.openms.de> (accessed on 23 February 2021)) software network (Figure 6a) [140].

Following the successful application of the workflow and encouraging results obtained in case of UV cross-linking of DNA-binding proteins in native chromatin, Stützer et al. [140], eventually applied it for analyzing protein–DNA interactions in more complex samples. Towards these ends, intact nuclei were isolated from HeLa cells and subjected to UV irradiation. Chromatin was isolated from the cross-linked nuclei by formaldehyde induced the cross-linking method based on chromatin-precipitation [148]. The isolated UV-irradiated chromatin fraction was digested while using RNase, DNase, and trypsin. The resulting mixture containing peptides, cross-linked species and oligonucleotides was then further processed (Figure 6a) [140]. In these initial in nucleosome cross-linking experiments, Stützer et al. [140] hardly observed histone cross-links, and they referred it to the used purification strategy (TiO₂-based enrichment without any further steps to remove the multitude of phosphorylated peptides or cross-links (peptide–RNA) that are present in UV cross-linked and digested nuclei). The co-enriched peptides block mass spectrometry detection as well as the sequencing of DNA cross-linked peptides. In order to overcome this obstacle, Stützer et al. [140] applied a more sophisticated enrichment strategy, in which they combined a size-exclusion chromatography (SEC) step along with chromatin-isolation and final TiO₂ affinity enrichment. By removing most RNAs by RNase digestion and trypsinizing protein and maintaining the DNA intact, the larger-sized DNA–peptide cross-links were successfully separated from linear (phosphorylated) peptides as well as from the vast majority of peptide–RNA oligonucleotide cross-links (Figure 6a).

2.5. Protein–RNA Interactions Identified by Cross-Linking MS Technique

Non-coding RNA sequences involving long non-coding RNAs, small nucleolar RNAs, and untranslated mRNA regions make direct interactions with proteins to achieve their various functions. Recent efforts have categorized the methods for studying RNA–protein interactions in two different approaches, one that identifies proteins that are bound to RNA of interest (RNA-centric) and other that characterizes RNAs bound to a protein of interest (protein-centric). Herein, we review different methods for studying protein–RNA interactions, while focusing on the cross-linking MS technique (Figure 6b) [149].

2.5.1. The RNA-Centric Cross-Linking

The changing patterns of RNA–protein interactions is critical for cellular functions, which are then regenerated based on the subcellular localization and environment stimuli [150]. Characterizing such protein–RNA interactions is challenging, as they are dynamic and transient. Generally, RNA-centric methods can be *in vitro* or *in vivo*, from which *in vitro* approaches are useful for analyzing RNA and protein molecules outside the context of a cell, while *in vivo* methods are useful to investigate such interactions within the cellular environment [149].

Cross-linking methods can be used to identify protein–RNA interactions *in vivo*. To do so, the RNA is purified under denaturing conditions to remove non-covalent interactions and only cross-linked components are subsequently extracted for identification. The formaldehyde, a small and bifunctional cross-linker, can easily permeate the cells and cross-links macromolecules within 2 Å, involving the protein–RNA complexes, by creating a reversible covalent linkage [151]. Methods that use formaldehyde to cross-link RNA to proteins are referred to as Chromatin Isolation by RNA Purification (ChIRP) [152] and capture hybridization analysis of RNA targets (CHART) (Figure 6b) [153]. Additionally, the UV light is a zero-distance cross-linker method and, thus, it cross-links protein to

nucleic acid at a zero distance or in direct contact and in irreversible covalent bonds. Despite that the UV light is considered to be a more specific cross-linker, the efficiency of UV cross-linking is lower, it has slight uridine preference, and double-stranded RNA is known to be poorly cross-linked [154]. In vivo methods that use UV cross-linking involve RNA affinity purification (RAP) [155], peptide-nucleic acid assisted identification of RNA-binding proteins (PAIR) [156], MS2 in vivo biotin-tagged RAP (MS2-BioTRAP) [157], and tandem RNA isolation procedure (TRIP) [158]. Although all of these methods use a UV cross-linking approach, they have a different experimental setup.

The strength of the RNA–protein interaction should be mainly considered while choosing the cross-linking approach. RNA–protein dissociation constants vary widely, and such experimentally measured constants can be in the range of high nanomolar to picomolar concentrations. Generally, the weaker RNA–protein interactions are less likely to be captured by UV cross-linking as contrary to formaldehyde cross-linking. In addition, the UV cross-linking efficiency varies according to amino acid chemistry, whereas, in the formaldehyde cross-linking, nucleophilic lysine residues are strongly preferred and cross-linked (Figure 6b) [149]. Hence, more cells can be required to capture RNA–protein interactions with UV light cross-linking as compared to that of the formaldehyde method [149].

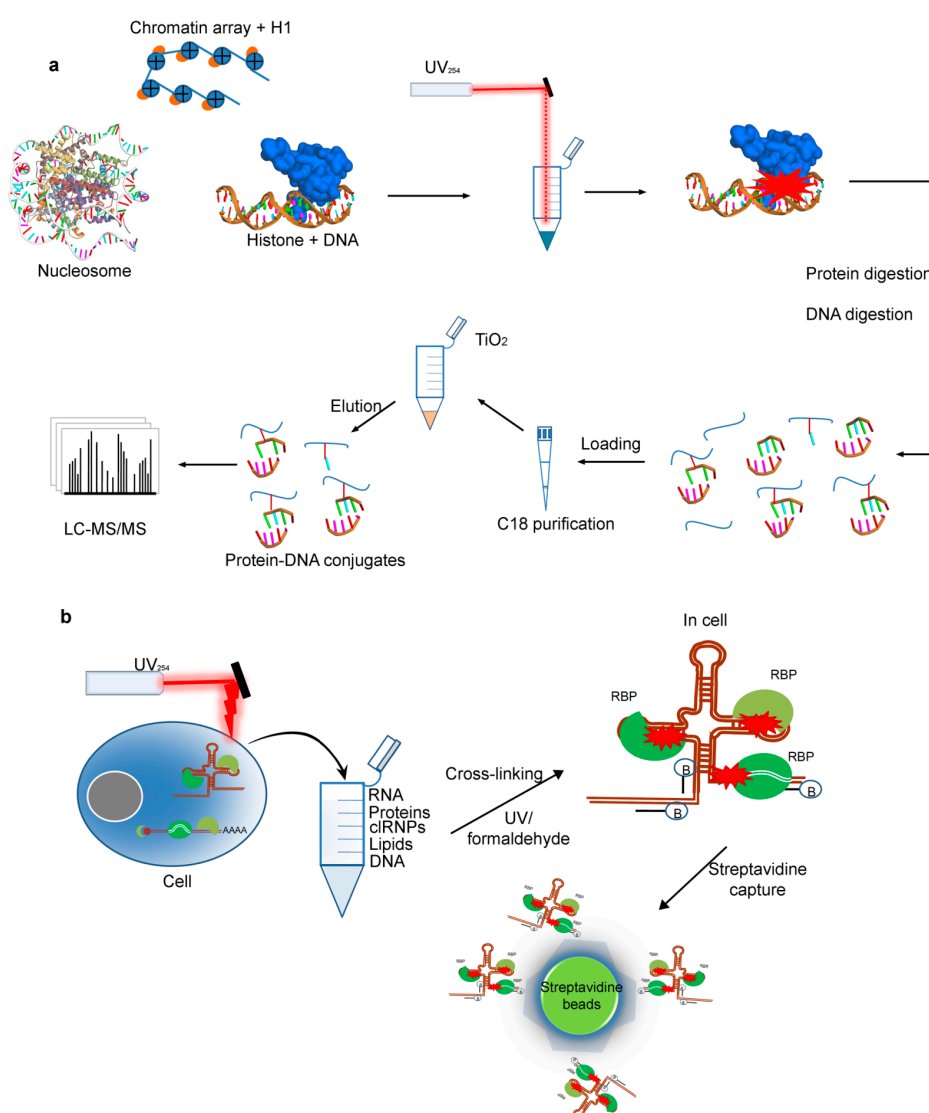


Figure 6. Schematic outline of the CLMS technique, to probe protein- DNA /RNA interactions. (a) Workflow for mass spectrometry analysis of UV-induced cross-links in histone-DNA complexes, nucleosomes, and chromatin arrays [140,159].

(b) Cross-linking based methods use either UV (RAP, RNA affinity purification; PAIR, peptide-nucleic acid assisted identification of RNA-binding proteins; MS2-BioTRAP, MS2 in vivo biotin-tagged RAP; and TRIP, tandem RNA isolation procedure) or formaldehyde cross-linking (CHART, capture hybridization analysis of RNA targets; ChIRP, Chromatin Isolation by RNA Purification) to capture RNA–protein interactions. Biotinylated oligonucleotide probes are hybridized to the RNA, and then the RNA and cross-linked proteins are purified for downstream analysis (RBP, RNA-binding protein; clRNPs, cross-linked RNPs) [149,160]. Representations of the protein or peptide structures in the figure, were prepared using BIOVIA Discovery Studio (Dassault Systèmes, BIOVIA Corp., San Diego, CA, USA) visualizer.

In the proteomic analysis, most of the RNA-centric methods follow quantitative mass spectrometry workflows to detect proteins that are bound to RNA. These MS approaches can be either label-free or incorporate the use of chemical labels for detection or simply quantitation. Labeling methods can use stable isotope labels or chemical tagging of proteins in samples (can be termed as controls) and, thus, the enrichment scores can be obtained by the ratio of labeled peptides and true binding partners can be identified [161]. Labeling MS techniques, like SILAC and iTRAQ, are especially effective with formaldehyde cross-linking. However, they are more expensive and desire greater technical expertise in MS data analysis. The main challenge of label-free MS is to distinguish true binding partners from non-specific proteins [162]. Analytical tools, for instance, SAINT (significance analysis of interactome score), can be useful for spectral count data from non-quantitative MS in order to effectively score the probability of a protein–RNA interaction in this case. Some of the methods require the purification of cellular RNA (e.g., CHART, RAP), validating the isolated RNA by sequencing will assure that proteomic analysis has certainly found interactions with RNA [149].

2.5.2. The Protein-Centric Cross-Linking

The majority of studies identify RNAs that are bound to a protein are performed by purifying the protein of interest. The most common way to achieve this is by making use of the long-known fact that the protein will chemically cross-link to the nucleic acid in vivo when hit by UV light at ~254 nm [163]. This UV induced cross-linking played a key role in identifying RNA binding proteins, and nearly all amino acids cross-links excluding the residues: aspartic acid, glutamic acid, asparagine, and glutamine [141]. Techniques in which the UV cross-linking is followed by the protein purification and bound RNAs can be identified are widely termed cross-linking immunoprecipitation (CLIP) methods [164], along with those using the high-throughput sequencing (HTS) forming the CLIP-seq family of methods [165]. Approaches using a similar protocol, but with an alternative cross-linker, are also specified as CLIP methods. In some cases, where the indirect interactions are not bearable, alternative cross-linking reagents can be effective, for example, the PAR-CLIP (photoactivatable-ribonucleoside-enhanced cross-linking and immunoprecipitation) [166]. PAR-CLIP uses 4-thiouridine and/or 5-thioguanine as a nucleotide analog, and it is beneficial, especially where the UV light is not penetrating deep enough into the sample. However, the results have been mostly similar to those with the regular cross-linking technique [149].

Lately, the formaldehyde has been used as a cross-linking reagent for CLIP method on a double-stranded RNA-binding protein, which are, in most cases, thought to UV cross-link poorly [166]. Methylene blue was used to cross-link double stranded RNA to RNA binding proteins [167,168]; however, this method has not been broadly used in CLIP so far. Many compounds that are known to cross-link RNA to protein, like Dithiothreitol [169], 2-iminothiolane [170], and diepoxybutane [171], are not considered to be reagents for CLIP methods and, likewise, many other cross-linking compounds remained uncharacterized. Despite the use of either standard UV cross-linking or alternative methods, approaches that are based on protein purification for protein-centric RNA studies establish a backbone of the field and, among these, the leading methods are the rapidly expanding array of CLIP [149].

Furthermore, to develop the mechanisms for analyzing protein–RNA interactions is one of the possible, but not really touched, potential of cross-linking reactions. In order to

identify such interactions, the use of diglyoxal compounds has been described in analyzing the ribosome organization [103,172]. These cross-linking reagents are determined to be applicable for studying protein structure with the identification of arginine-arginine cross-links [173] in proteins, as well as also useful in analyzing the nucleic acid structure [103,174]. AN Holding has described the successful use of diglyoxal cross-linkers [103,174], by developing a ^{13}C labeled form of biphenyldiglyoxal (Figure 4e) from the friedel-crafts acylation reaction between biphenyl and $^{13}\text{C}_2$ -acetyl chloride, and that is then followed by oxidation of the terminal carbons using HBr/DMSO oxidation.

2.6. Pairing the CLMS Methodologies with Molecular Dynamics Simulations

Over the years, there has been a collection of biological data recorded for the purpose of building virtual biological models that can also be used for molecular dynamics simulation (MDS). Based on a general model of the physics (or biophysics), the MD simulation predicts the conformational dynamics or time course movements (at femtosecond resolution) of every atom from a biomolecule (e.g., protein, RNA, DNA, etc.) that is used to assist understanding of important biomolecular processes [175–178]. These simulation methodologies are often applied alongside several experimental techniques, for example, NMR [179,180]. The NMR technique can efficiently resolve secondary structure; however, a pairing between NMR cross-coefficients and MDS under NMR restraints/constraints enables the construction of self-consistent complete protein tertiary structures, which could even resemble physiological tertiary structures. These unique features of the deterministic MD simulation technique suggest that the MS-based approaches, such as the chemical-cross linking in combination with MDS, could play an important role in structural biology, and bring insights for several biomolecular processes. The CLMS methodologies provide inter-residue distances that can be integrated into the molecular modeling and simulation techniques (especially coarse grained MD) to achieve physiologically realistic quaternary (PPI) structures, which can be difficult to resolve by techniques, such as X-ray crystallography or cryo-EM.

Currently, significant progress is being made by pairing the cross-linking mass spectrometry with MDS in order to explore a wide range of biological questions concerning protein motions, interactions, and their assemblies. For example, Brodie et al. presented an integrative structural biology approach, in which short-distance cross-linking constraints are incorporated into rapid discrete molecular dynamics (DMD) simulations [181]. They provided a workflow on proteins with well-defined structures, and they have also validated the predicted structural models with other experimental structural proteomics approaches, namely: hydrogen-deuterium exchange, chemical surface modification, and long-distance cross-linking. Three main steps of their workflow involve: (i) the acquisition of short-distance cross-linking data, (ii) performance of DMD simulations that are guided by these cross-linking constraints, and (iii) validation of the obtained structures with additional structural proteomics methods [181,182]. Moreover, despite the in-depth information derived from the CLMS techniques, the cross-linking experiments can occasionally generate inconsistent data due to the fluctuations in solution structures of protein [181,183] and, thus, the inclusion of cross-linking constraints will define a structural ensemble instead of a single protein structure. Therefore, one must consider this while selecting the best-fit models from computationally produced ensembles of conformations [181,184], as well as while directly integrating distance constraints into an energy-based simulation process [181,185].

3. Conclusions and Outlook

Rapidly growing technologies to map out interactions between protein–protein /RNA/DNA are critically important since the biological function of these molecules is extremely influenced by their structures, complex formation, locations and regulatory networks. Herein, we reviewed a suite of methods that are valuable for detecting such interactions between biomolecules of the cell. Special emphasis was on the combination of cross-linking with native mass spectrometry technique to yield mutual benefits while

characterizing protein–protein, protein–RNA, or protein–DNA interactions, though, with a slightly modified approach. The CLMS strategies allow for capturing and identifying not just stable, but also transient, dynamic, and weakly bonding molecules and, thus, it emerged as the most striking example of multidisciplinary success among hybrid or integrative structural biology methods. In addition, CLMS provides inter-residue distances that can be integrated into the molecular modeling and MD simulation techniques, in order to achieve physiologically realistic quaternary (PPI) structures. Different means to combine the knowledge of chemistry with MS to analyze essential biological systems of interest were scrutinized. Several different cross-linkers are accessible with varied chemistries for the CLMS techniques and, therefore, one should consider the protein sequence of interest to determine what sort of combination of cross-linker and cleavage enzyme yield positive results. To start with one of the most commonly used cross-linker types, like amine reactive cross-linking (NH₂–NH₂: BS3 and DSS, spacer length 11.4 Å; and BS2G, spacer length 7.7 Å) and carboxyl-amine cross-linking (COOH–NH₂: EDC or sulfo-NHS zero spacer length), is advisable.

Furthermore, we propose (or speculate) that the CLMS approach, though, not previously carried out, could be merged with novel interactomics techniques (proximity-dependent labeling by BioID or TurboID). Merging these two methods may bring spatial resolution to CLMS at a sub organelle level, because the BioID radius is estimated ~10 nm, and since CLMS mainly focuses on the proteins in the neighborhood of a given bait. These CLMS and proximal interactomics methods can be iteratively performed in living cells, i.e., the generated samples are the proximal species of a protein of interest labeled with biotin first and then cross-linked with interactors. Subsequently, perform streptavidin-biotin capture, followed by the digestion and cross-linked peptides identification of proximal interactors.

Besides the outstanding practice of CLMS in cross-linked peptide identification, it is worth noting that the failure to detect cross-links or only a few cross-links is related to several factors. For example, the lengths of the cross-linked peptides are too short or long, defective fragmentation of the cross-linked peptide, the cross-linked peptide is below the MS detection threshold, or the cross-linker is too short to link the appropriate residues. A few additional factors are: the sequence of interest is not fitting well with the cross-linker chemistry concerning reactive or cleavable residues, imperfect reaction conditions, the whole amount is too low, and extraction difficulties of the cross-linked peptides from a gel (in such a case, ‘in-solution’ digest is an alternative option). Therefore, the functional groups in a linker have to be tested and characterized thoroughly, as well as demonstrated precisely within the CLMS framework. Additionally, we addressed different protocols that are useful for digestion in CLMS methodologies, such as ‘in-gel’, ‘out-gel’, ‘in solution’, and FASP, with their successful usage. Alongside the CLMS application for identifying the protein–protein interactions, this review covers or describes the cross-linking methods for studying protein–RNA as well as protein–DNA interactions, which apparently arises as fertile ground for future CLMS utilization.

Author Contributions: Conceptualization, U.K., M.P., J.F., D.R.G. and T.R.H.; formal analysis, U.K., M.P., J.F. and D.R.G.; investigation, U.K., M.P. and J.F.; resources, U.K. and T.R.H.; data curation, U.K., M.P., J.F., D.R.G. and T.R.H.; writing—original draft preparation, U.K., M.P., J.F., R.F., D.R.G. and T.R.H.; writing—review and editing, U.K., M.P., J.F., E.C., J.A.A., R.F., D.R.G. and T.R.H.; visualization, U.K. and M.P.; project administration, U.K., D.R.G. and T.R.H. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the International Centre for Cancer Vaccine Science, University of Gdansk (Fundacja na rzecz Nauki Polskiej: MAB/3/2017). U.K. is supported by the grant: 2020/36/C/NZ2/00108, from The National Science Centre (Narodowe Centrum Nauki, Krakow, Poland).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The International Centre for Cancer Vaccine Science project is carried out within the International Research Agendas programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund. The University of Victoria-Genome BC Proteomics Centre is grateful to Genome Canada and Genome British Columbia for financial support for Genomics Technology Platforms (GTP) funding for operations and technology development (264PRO).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhou, M.; Li, Q.; Wang, R. Current experimental methods for characterizing protein-protein interactions. *Chem. Med. Chem.* **2016**, *11*, 738–756. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Berggård, T.; Linse, S.; James, P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* **2007**, *16*, 2833–2842. [\[CrossRef\]](#)
3. Brown, C.J.; Lain, S.; Verma, C.S.; Fersht, A.R.; Lane, D.P. Awakening guardian angels: Drugging the p53 pathway. *Nat. Rev. Cancer* **2009**, *12*, 862–873. [\[CrossRef\]](#)
4. Joerger, A.C.; Fersht, A.R. Structure-function-rescue: The diverse nature of common p53 cancer mutants. *Oncogene* **2007**, *26*, 2226–2242. [\[CrossRef\]](#)
5. Brooke, M.A.; Nitoiu, D.; Kelsell, D.P. Cell-cell connectivity: Desmosomes and disease. *J. Pathol.* **2012**, *226*, 158–171. [\[CrossRef\]](#)
6. Heck, A.J. Native mass spectrometry: A bridge between interactomics and structural biology. *Nat. Methods* **2008**, *5*, 927–933. [\[CrossRef\]](#)
7. Pirrone, G.F.; Iacob, R.E.; Engen, J.R. Applications of hydrogen/deuterium exchange MS from 2012 to 2014. *Anal. Chem.* **2015**, *87*, 99–118. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Tugarinov, V.; Hwang, P.M.; Kay, L.E. Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev. Biochem.* **2004**, *73*, 107–146. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Ilari, A.; Savino, C. Protein structure determination by x-ray crystallography. *Methods Mol. Biol.* **2008**, *452*, 63–87. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Walian, P.; Cross, T.A.; Jap, B.K. Structural genomics of membrane proteins. *Genome Biol.* **2004**, *5*, 215. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Iacobucci, C.; Götze, M.; Sinz, A. Cross-linking/mass spectrometry to get a closer view on protein interaction networks. *Curr. Opin. Biotechnol.* **2020**, *63*, 48–53. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nat. Methods* **2005**, *2*, 201–206. [\[CrossRef\]](#)
13. Farrell, I.S.; Toroney, R.; Hazen, J.L.; Mehl, R.A.; Chin, J.W. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat. Methods* **2005**, *2*, 377–384. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Milne, J.L.; Borgnia, M.J.; Bartesaghi, A.; Tran, E.E.; Earl, L.A.; Schauder, D.M.; Lengyel, J.; Pierson, J.; Patwardhan, A.; Subramaniam, S. Cryo-electron microscopy—a primer for the nonmicroscopist. *FEBS J.* **2013**, *280*, 28–45. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Hura, G.L.; Menon, A.L.; Hammel, M.; Rambo, R.P.; Poole, F.L., II; Tsutakawa, S.E.; Jenney, F.E., Jr.; Classen, S.; Frankel, K.A.; Hopkins, R.C.; et al. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Methods* **2009**, *6*, 606–612. [\[CrossRef\]](#)
16. Uetrecht, C.; Rose, R.J.; Van Duijn, E.; Lorenzen, K.; Heck, A.J. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem Soc. Rev.* **2010**, *39*, 1633–1655. [\[CrossRef\]](#)
17. Konermann, L.; Pan, J.; Liu, Y.H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **2011**, *40*, 1224–1234. [\[CrossRef\]](#)
18. Gau, B.C.; Sharp, J.S.; Rempel, D.L.; Gross, M.L. Fast photochemical oxidation of protein footprints faster than protein unfolding. *Anal. Chem.* **2009**, *81*, 6563–6571. [\[CrossRef\]](#)
19. Perrin, F. Polarisation de fluorescence and mean life of excited molecules. Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l'état excité. *J. Phys. Radium.* **1926**, *7*, 390–401.
20. Du, Y. *Protein-Protein Interactions, of the Series Methods in Molecular Biology*; Meyerkord, C.L., Fu, H., Walker, J.M., Eds.; Springer: New York, NY, USA, 2015; Volume 1278, pp. 529–544.
21. Madeira, A.; Ohman, E.; Nilsson, A.; Sjögren, B.; Andrén, P.E.; Svenningsson, P. Coupling surface plasmon resonance to mass spectrometry to discover novel protein-protein interactions. *Nat. Protoc.* **2009**, *4*, 1023–1037. [\[CrossRef\]](#)
22. Boozer, C.; Kim, G.; Cong, S.; Guan, H.; Londergan, T. Looking towards label-free biomolecular interaction analysis in a high-throughput format: A review of new surface plasmon resonance technologies. *Curr. Opin. Biotechnol.* **2006**, *17*, 400–405. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Thompson, P.M.; Beck, M.R.; Campbell, S.L. *Protein-Protein Interactions, of the Series Methods in Molecular*; Springer: New York, NY, USA, 2015; Volume 1278, pp. 267–279.
24. Ranjbar, B.; Gill, P. Circular dichroism techniques: Biomolecular and nanostructural analyses—A review. *Chem. Biol. Drug. Des.* **2009**, *74*, 101–120. [\[CrossRef\]](#) [\[PubMed\]](#)

25. Some, D. Light-scattering-based analysis of biomolecular interactions. *Biophys. Rev.* **2013**, *5*, 147–158. [[CrossRef](#)] [[PubMed](#)]
26. Laue, T. Analytical ultracentrifugation: A powerful ‘new’ technology in drug discovery. *Drug Discov. Today Technol.* **2004**, *1*, 309–315. [[CrossRef](#)] [[PubMed](#)]
27. Velazquez-Campoy, A.; Leavitt, S.A.; Freire, E. Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods Mol. Biol.* **2015**, 1278, 183–204. [[CrossRef](#)]
28. Wienken, C.J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **2010**, *1*, 100. [[CrossRef](#)]
29. Selvin, P.R. The renaissance of fluorescence resonance energy transfer. *Nat. Struct. Biol.* **2000**, *7*, 730–734. [[CrossRef](#)]
30. Pflieger, K.D.; Eidne, K.A. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). *Nat. Methods* **2006**, *3*, 165–174. [[CrossRef](#)]
31. Yasgar, A.; Jadhav, A.; Simeonov, A.; Coussens, N.P. AlphaScreen-based assays: Ultra-high-throughput screening for small-molecule inhibitors of challenging enzymes and protein-protein interactions. *Methods Mol. Biol.* **2016**, 1439, 77–98. [[CrossRef](#)]
32. Remy, I.; Campbell-Valois, F.X.; Michnick, S.W. Detection of protein-protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase. *Nat. Protoc.* **2007**, *2*, 2120–2125. [[CrossRef](#)]
33. Beeckmans, S. Chromatographic methods to study protein-protein interactions. *Methods* **1999**, *19*, 278–305. [[CrossRef](#)]
34. Muronetz, V.I.; Sholukh, M.; Korpela, T. Use of protein-protein interactions in affinity chromatography. *J. Biochem. Biophys. Methods* **2001**, *49*, 29–47. [[CrossRef](#)]
35. Smith, G.P. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228*, 1315–1317. [[CrossRef](#)]
36. Pande, J.; Szewczyk, M.M.; Grover, A.K. Phage display: Concept, innovations, applications and future. *Biotechnol. Adv.* **2010**, *28*, 849–858. [[CrossRef](#)]
37. Fields, S.; Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **1989**, *340*, 245–246. [[CrossRef](#)] [[PubMed](#)]
38. Hamdi, A.; Colas, P. Yeast two-hybrid methods and their applications in drug discovery. *Trends Pharmacol. Sci.* **2012**, *33*, 109–118. [[CrossRef](#)] [[PubMed](#)]
39. Hall, D.A.; Ptacek, J.; Snyder, M. Protein microarray technology. *Mech. Ageing Dev.* **2007**, *128*, 161–167. [[CrossRef](#)] [[PubMed](#)]
40. Zhu, H.; Bilgin, M.; Bangham, R.; Hall, D.; Casamayor, A.; Bertone, P.; Lan, N.; Jansen, R.; Bidlingmaier, S.; Houfek, T.; et al. Global analysis of protein activities using proteome chips. *Science* **2001**, *293*, 2101–2105. [[CrossRef](#)] [[PubMed](#)]
41. Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64–71. [[CrossRef](#)] [[PubMed](#)]
42. Karas, M.; Hillenkamp, F. Laser desorption/ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988**, *60*, 2299–2301. [[CrossRef](#)]
43. Bich, C.; Zenobi, R. Mass spectrometry of large complexes. *Curr. Opin. Struct. Biol.* **2009**, *19*, 632–639. [[CrossRef](#)]
44. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151–153. [[CrossRef](#)]
45. Glush, G.L.; Vachet, R.W. The basics of mass spectrometry in the twenty-first century. *Nat. Rev. Drug Discov.* **2003**, *2*, 140–150. [[CrossRef](#)] [[PubMed](#)]
46. Kukacka, Z.; Rosulek, M.; Strohal, M.; Kavan, D.; Novak, P. Mapping protein structural changes by quantitative cross-linking. *Methods* **2015**, *89*, 112–120. [[CrossRef](#)] [[PubMed](#)]
47. Fischer, L.; Chen, Z.A.; Rappsilber, J. Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J. Proteom.* **2013**, *88*, 120–128. [[CrossRef](#)]
48. Ong, S.E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D.B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteom.* **2002**, *1*, 376–386. [[CrossRef](#)]
49. Chavez, J.D.; Schweppe, D.K.; Eng, J.K.; Bruce, J.E. In Vivo Conformational Dynamics of Hsp90 and Its Interactors. *Cell Chem. Biol.* **2016**, *23*, 716–726. [[CrossRef](#)]
50. Pearson, K.M.; Pannell, L.K.; Fales, H.M. Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 149–159. [[CrossRef](#)]
51. Müller, F.; Graziadei, A.; Rappsilber, J. Quantitative Photo-crosslinking Mass Spectrometry Revealing Protein Structure Response to Environmental Changes. *Anal. Chem.* **2019**, *91*, 9041–9048. [[CrossRef](#)]
52. Chen, Z.A.; Rappsilber, J. Quantitative cross-linking/mass spectrometry to elucidate structural changes in proteins and their complexes. *Nat. Protoc.* **2019**, *14*, 171–201. [[CrossRef](#)] [[PubMed](#)]
53. Chen, Z.; Fischer, L.; Tahir, S.; Bukowski-Wills, J.C.; Barlow, P.; Rappsilber, J. Quantitative cross-linking/mass spectrometry reveals subtle protein conformational changes. *Wellcome Open Res.* **2016**, *1*, 5. [[CrossRef](#)]
54. Chen, Z.A.; Rappsilber, J. Protein Dynamics in Solution by Quantitative Crosslinking/Mass Spectrometry. *Trends Biochem. Sci.* **2018**, *43*, 908–920. [[CrossRef](#)] [[PubMed](#)]
55. Kleiner, R.E.; Hang, L.E.; Molloy, K.R.; Chait, B.T.; Kapoor, T.M. A Chemical Proteomics Approach to Reveal Direct Protein-Protein Interactions in Living Cells. *Cell Chem. Biol.* **2018**, *25*, 110–120.e3. [[CrossRef](#)]

56. Yu, C.; Huszagh, A.; Viner, R.; Novitsky, E.J.; Rychnovsky, S.D.; Huang, L. Developing a Multiplexed Quantitative Cross-Linking Mass Spectrometry Platform for Comparative Structural Analysis of Protein Complexes. *Anal. Chem.* **2016**, *88*, 10301–10308. [[CrossRef](#)] [[PubMed](#)]
57. Müller, F.; Rappsilber, J. A protocol for studying structural dynamics of proteins by quantitative crosslinking mass spectrometry and data-independent acquisition. *J. Proteom.* **2020**, *218*, 103721. [[CrossRef](#)]
58. Zhou, X.Z.; Lu, K.P. The isomerase PIN1 controls numerous cancer-driving pathways and is a unique drug target. *Nat. Rev. Cancer* **2016**, *16*, 463–478. [[CrossRef](#)]
59. Shaw, P.E. Peptidyl-prolyl isomerases: A new twist to transcription. *EMBO Rep.* **2002**, *3*, 521–526. [[CrossRef](#)]
60. Lin, W.; Bonin, M.; Boden, A.; Wieduwild, R.; Murawala, P.; Wermke, M.; Andrade, H.; Bornhäuser, M.; Zhang, Y. Peptidyl prolyl cis/trans isomerase activity on the cell surface correlates with extracellular matrix development. *Commun. Biol.* **2019**, *2*, 58. [[CrossRef](#)] [[PubMed](#)]
61. Chavez, J.D.; Bruce, J.E. Chemical cross-linking with mass spectrometry: A tool for systems structural biology. *Curr. Opin. Chem. Biol.* **2019**, *48*, 8–18. [[CrossRef](#)]
62. Keller, A.; Chavez, J.D.; Tang, X.; Bruce, J.E. Leveraging the entirety of the protein data bank to enable improved structure prediction based on cross-link data. *J. Proteome Res.* **2021**, *20*, 1087–1095. [[CrossRef](#)]
63. Chavez, J.D.; Mohr, J.P.; Mathay, M.; Zhong, X.; Keller, A.; Bruce, J.E. Systems structural biology measurements by in vivo cross-linking with mass spectrometry. *Nat. Protoc.* **2019**, *14*, 2318–2343. [[CrossRef](#)]
64. Keller, A.; Chavez, J.D.; Felt, K.C.; Bruce, J.E. Prediction of an upper limit for the fraction of interprotein cross-links in large-scale in vivo cross-linking studies. *J. Proteome Res.* **2019**, *18*, 3077–3085. [[CrossRef](#)] [[PubMed](#)]
65. Liu, F.; Lössl, P.; Scheltema, R.; Viner, R.; Heck, A. Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nat. Commun.* **2017**, *8*, 15473. [[CrossRef](#)]
66. Klykov, O.; Steigenberger, B.; Pektaş, S.; Fasci, D.; Heck, A.; Scheltema, R.A. Efficient and robust proteome-wide approaches for cross-linking mass spectrometry. *Nat. Protoc.* **2018**, *13*, 2964–2990. [[CrossRef](#)]
67. Demartino, G.N. Reconstitution of PA700, the 19S regulatory particle, from purified precursor complexes. *Methods Mol. Biol.* **2012**, *832*, 443–452. [[CrossRef](#)] [[PubMed](#)]
68. Oeffinger, M. Two steps forward-one step back: Advances in affinity purification mass spectrometry of macromolecular complexes. *Proteomics* **2012**, *12*, 1591–1608. [[CrossRef](#)]
69. Mishra, P.K.; Yoo, C.M.; Hong, E.; Rhee, H.W. Photo-crosslinking: An emerging chemical tool for investigating molecular networks in live cells. *ChemBioChem* **2020**, *21*, 924–932. [[CrossRef](#)]
70. Tang, X.; Bruce, J.E. *Mass Spectrometry of Proteins and Peptides, of the Series Methods in Molecular*; Lipton, M.S., Paša-Tolic, L., Walker, J.M., Eds.; Humana Press: Totowa, NJ, USA, 2009; Volume 492, pp. 283–293.
71. Guerrero, C.; Tagwerker, C.; Kaiser, P.; Huang, L. An integrated mass spectrometry-based proteomic approach: Quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol. Cell Proteom.* **2006**, *5*, 366–378. [[CrossRef](#)] [[PubMed](#)]
72. Vasilescu, J.; Guo, X.; Kast, J. Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry. *Proteomics* **2004**, *4*, 3845–3854. [[CrossRef](#)]
73. Van den Heuvel, R.H.; Heck, A.J. Native protein mass spectrometry: From intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* **2004**, *8*, 519–526. [[CrossRef](#)] [[PubMed](#)]
74. Benesch, J.L.; Robinson, C.V. Mass spectrometry of macromolecular assemblies: Preservation and dissociation. *Curr. Opin. Struct. Biol.* **2006**, *16*, 245–251. [[CrossRef](#)]
75. Sharon, M.; Robinson, C.V. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annu. Rev. Biochem.* **2007**, *76*, 167–193. [[CrossRef](#)]
76. Synowsky, S.A.; Van den Heuvel, R.H.; Mohammed, S.; Pijnappel, P.W.; Heck, A.J. Probing genuine strong interactions and post-translational modifications in the heterogeneous yeast exosome protein complex. *Mol. Cell Proteom.* **2006**, *5*, 1581–1592. [[CrossRef](#)] [[PubMed](#)]
77. Yu, C.; Huang, L. Cross-linking mass spectrometry: An emerging technology for interactomics and structural biology. *Anal. Chem.* **2018**, *90*, 144–165. [[CrossRef](#)]
78. Sinz, A. Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal. Bioanal. Chem.* **2010**, *397*, 3433–3440. [[CrossRef](#)]
79. Bruce, J.E. In vivo protein complex topologies: Sights through a cross-linking lens. *Proteomics* **2012**, *12*, 1565–1575. [[CrossRef](#)]
80. Piersimoni, L.; Sinz, A. Cross-linking/mass spectrometry at the crossroads. *Anal. Bioanal. Chem.* **2020**, *412*, 5981–5987. [[CrossRef](#)]
81. Leitner, A. Cross-linking and other structural proteomics techniques: How chemistry is enabling mass spectrometry applications in structural biology. *Chem. Sci.* **2016**, *7*, 4792–4803. [[CrossRef](#)] [[PubMed](#)]
82. Leitner, A.; Bonvin, A.M.J.J.; Borchers, C.H.; Chalkley, R.J.; Chamot-Rooke, J.; Combe, C.W.; Cox, J.; Dong, M.Q.; Fischer, L.; Götze, M.; et al. Toward increased reliability, transparency, and accessibility in cross-linking mass spectrometry. *Structure* **2020**, *28*, 1259–1268. [[CrossRef](#)] [[PubMed](#)]
83. Kao, A.; Chiu, C.L.; Vellucci, D.; Yang, Y.; Patel, V.R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S.D.; Huang, L. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell Proteom.* **2011**, *10*, M110.002212. [[CrossRef](#)]

84. Leitner, A.; Joachimiak, L.A.; Unverdorben, P.; Walzthoeni, T.; Frydman, J.; Förster, F.; Aebersold, R. Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9455–9460. [\[CrossRef\]](#)
85. Pinard, R.; Lambert, D.; Heckman, J.E.; Esteban, J.A.; Gundlach, C.W., 4th; Hampel, K.J.; Glick, G.D.; Walter, N.G.; Major, F.; Burke, J.M. The hairpin ribozyme substrate binding-domain: A highly constrained D-shaped conformation. *J. Mol. Biol.* **2001**, *307*, 51–65. [\[CrossRef\]](#)
86. Harris, M.E.; Christian, E.L. RNA crosslinking methods. *Methods Enzymol.* **2009**, *468*, 127–146. [\[CrossRef\]](#)
87. Buxbaum, E. Cross-linkers. In *Biophysical Chemistry of Proteins*; Springer: Boston, MA, USA, 2011.
88. Chen, F.; Nielsen, S.; Zenobi, R. Understanding chemical reactivity for homo- and heterobifunctional protein cross-linking agents. *J. Mass Spectrom.* **2013**, *48*, 807–812. [\[CrossRef\]](#) [\[PubMed\]](#)
89. Belsom, A.; Rappsilber, J. Anatomy of a crosslinker. *Curr. Opin. Chem. Biol.* **2020**, *60*, 39–46. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Steigenberger, B.; Albanese, P.; Heck, A.J.R.; Scheltema, R.A. To cleave or not to cleave in XL-MS? *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 196–206. [\[CrossRef\]](#)
91. Iacobucci, C.; Piotrowski, C.; Aebersold, R.; Amaral, B.C.; Andrews, P.; Bernfur, K.; Borchers, C.; Brodie, N.I.; Bruce, J.E.; Cao, Y.; et al. First community-wide, comparative cross-linking mass spectrometry study. *Anal. Chem.* **2019**, *91*, 6953–6961. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Trester-Zedlitz, M.; Kamada, K.; Burley, S.K.; Fenyő, D.; Chait, B.T.; Muir, T.W. A modular cross-linking approach for exploring protein interactions. *J. Am. Chem. Soc.* **2003**, *125*, 2416–2425. [\[CrossRef\]](#)
93. Tang, X.; Bruce, J.E. A new cross-linking strategy: Protein interaction reporter (PIR) technology for protein-protein interaction studies. *Mol. Biosyst.* **2010**, *6*, 939–947. [\[CrossRef\]](#) [\[PubMed\]](#)
94. Soderblom, E.J.; Goshe, M.B. Collision-induced dissociative chemical cross-linking reagents and methodology: Applications to protein structural characterization using tandem mass spectrometry analysis. *Anal. Chem.* **2006**, *78*, 8059–8068. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Liu, F.; Wu, C.; Sweedler, J.V.; Goshe, M.B. An enhanced protein crosslink identification strategy using CID-cleavable chemical crosslinkers and LC/MS(n) analysis. *Proteomics* **2012**, *12*, 401–405. [\[CrossRef\]](#)
96. Müller, M.Q.; Dreier, F.; Ihling, C.H.; Schäfer, M.; Sinz, A. Cleavable cross-linker for protein structure analysis: Reliable identification of cross-linking products by tandem MS. *Anal. Chem.* **2010**, *82*, 6958–6968. [\[CrossRef\]](#)
97. Petrotchenko, E.V.; Serpa, J.J.; Borchers, C.H. An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell Proteom.* **2011**, *10*, M110.001420. [\[CrossRef\]](#)
98. Clifford-Nunn, B.; Showalter, H.D.; Andrews, P.C. Quaternary diamines as mass spectrometry cleavable crosslinkers for protein interactions. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 201–212. [\[CrossRef\]](#)
99. Burke, A.M.; Kandur, W.; Novitsky, E.J.; Kaake, R.M.; Yu, C.; Kao, A.; Vellucci, D.; Huang, L.; Rychnovsky, S.D. Synthesis of two new enrichable and MS-cleavable cross-linkers to define protein-protein interactions by mass spectrometry. *Org. Biomol. Chem.* **2015**, *13*, 5030–5037. [\[CrossRef\]](#)
100. Hagen, S.E.; Liu, K.; Jin, Y.; Piersimoni, L.; Andrews, P.C.; Showalter, H.D. Synthesis of CID-cleavable protein crosslinking agents containing quaternary amines for structural mass spectrometry. *Org. Biomol. Chem.* **2018**, *16*, 8245–8248. [\[CrossRef\]](#) [\[PubMed\]](#)
101. Steigenberger, B.; Pieters, R.J.; Heck, A.J.R.; Scheltema, R.A. PhoX: An imac-enrichable cross-linking reagent. *ACS Cent. Sci.* **2019**, *5*, 1514–1522. [\[CrossRef\]](#) [\[PubMed\]](#)
102. Chowdhury, S.M.; Du, X.; Tolić, N.; Wu, S.; Moore, R.J.; Mayer, M.U.; Smith, R.D.; Adkins, J.N. Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry. *Anal. Chem.* **2009**, *81*, 5524–5532. [\[CrossRef\]](#)
103. Holding, A.N. XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods San Diego Calif.* **2015**, *89*, 54–63. [\[CrossRef\]](#)
104. Turnbough, C.L., Jr.; Switzer, R.L. Regulation of pyrimidine biosynthetic gene expression in bacteria: Repression without repressors. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 266–300. [\[CrossRef\]](#) [\[PubMed\]](#)
105. 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*, 1193–1209. [\[CrossRef\]](#)
106. Götze, M.; Pettelkau, J.; Fritzsche, R.; Ihling, C.H.; Schäfer, M.; Sinz, A. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 83–97. [\[CrossRef\]](#)
107. Lima, D.B.; De Lima, T.B.; Balbuena, T.S.; Neves-Ferreira, A.; Barbosa, V.C.; Gozzo, F.C.; Carvalho, P.C. SIM-XL: A powerful and user-friendly tool for peptide cross-linking analysis. *J. Proteom.* **2015**, *129*, 51–55. [\[CrossRef\]](#)
108. Yılmaz, Ş.; Drepper, F.; Hulstaert, N.; Černič, M.; Gevaert, K.; Economou, A.; Warscheid, B.; Martens, L.; Vandermarliere, E. Xilmass: A New Approach toward the Identification of Cross-Linked Peptides. *Anal. Chem.* **2016**, *88*, 9949–9957. [\[CrossRef\]](#)
109. Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller, L.N.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R. Identification of cross-linked peptides from large sequence databases. *Nat. Methods* **2008**, *5*, 315–318. [\[CrossRef\]](#)
110. Mendes, M.L.; Fischer, L.; Chen, Z.A.; Barbon, M.; O'Reilly, F.J.; Giese, S.H.; Bohlke-Schneider, M.; Belsom, A.; Dau, T.; Combe, C.W.; et al. An integrated workflow for crosslinking mass spectrometry. *Mol. Syst. Biol.* **2019**, *15*, e8994. [\[CrossRef\]](#)
111. Hoopmann, M.R.; Zelter, A.; Johnson, R.S.; Riffle, M.; MacCoss, M.J.; Davis, T.N.; Moritz, R.L. Kojak: Efficient analysis of chemically cross-linked protein complexes. *J. Proteome Res.* **2015**, *14*, 2190–2198. [\[CrossRef\]](#) [\[PubMed\]](#)

112. Chen, Z.L.; Meng, J.M.; Cao, Y.; Yin, J.L.; Fang, R.Q.; Fan, S.B.; Liu, C.; Zeng, W.F.; Ding, Y.H.; Tan, D.; et al. A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat. Commun.* **2019**, *10*, 3404. [\[CrossRef\]](#)
113. Müller, F.; Kolbowski, L.; Bernhardt, O.M.; Reiter, L.; Rappsilber, J. Data-independent acquisition improves quantitative cross-linking mass spectrometry. *Mol. Cell Proteom.* **2019**, *18*, 786–795. [\[CrossRef\]](#) [\[PubMed\]](#)
114. Maiolica, A.; Cittaro, D.; Borsotti, D.; Sennels, L.; Ciferri, C.; Tarricone, C.; Musacchio, A.; Rappsilber, J. Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. *Mol. Cell Proteom.* **2007**, *6*, 2200–2211. [\[CrossRef\]](#) [\[PubMed\]](#)
115. Algret, R.; Fernandez-Martinez, J.; Shi, Y.; Kim, S.J.; Pellarin, R.; Cimerancic, P.; Cochet, E.; Sali, A.; Chait, B.T.; Rout, M.P.; et al. Molecular architecture and function of the SEA complex, a modulator of the TORC1 pathway. *Mol. Cell Proteom.* **2014**, *13*, 2855–2870. [\[CrossRef\]](#)
116. Pettelkau, J.; Thondorf, I.; Theisgen, S.; Lilie, H.; Schröder, T.; Arlt, C.; Ihling, C.H.; Sinz, A. Structural analysis of guanylyl cyclase-activating protein-2 (GCAP-2) homodimer by stable isotope-labeling, chemical cross-linking, and mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1969–1979. [\[CrossRef\]](#) [\[PubMed\]](#)
117. Arlt, C.; Ihling, C.H.; Sinz, A. Structure of full-length p53 tumor suppressor probed by chemical cross-linking and mass spectrometry. *Proteomics* **2015**, *15*, 2746–2755. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Hotta, K.; Kurokawa, M. Separation of fucose and acetylhexosamines by two-dimensional thin-layer chromatography. *Anal. Biochem.* **1968**, *26*, 472–474. [\[CrossRef\]](#)
119. Petrotchenko, E.V.; Serpa, J.J.; Cabecinha, A.N.; Lesperance, M.; Borchers, C.H. “Out-gel” tryptic digestion procedure for chemical cross-linking studies with mass spectrometric detection. *J. Proteome Res.* **2014**, *13*, 527–535. [\[CrossRef\]](#)
120. Parfentev, I.; Schilbach, S.; Cramer, P.; Urlaub, H. An experimentally generated peptide database increases the sensitivity of XL-MS with complex samples. *J. Proteom.* **2020**, *220*, 103754. [\[CrossRef\]](#)
121. 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*, 530–540. [\[CrossRef\]](#) [\[PubMed\]](#)
122. Mintseris, J.; Gygi, S.P. High-density chemical cross-linking for modeling protein interactions. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 93–102. [\[CrossRef\]](#) [\[PubMed\]](#)
123. Gonzalez-Lozano, M.A.; Koopmans, F.; Sullivan, P.F.; Protze, J.; Krause, G.; Verhage, M.; Li, K.W.; Liu, F.; Smit, A.B. Stitching the synapse: Cross-linking mass spectrometry into resolving synaptic protein interactions. *Sci. Adv.* **2020**, *6*, eaax5783. [\[CrossRef\]](#)
124. Slavin, M.; Kalisman, N. Structural analysis of protein complexes by cross-linking and mass spectrometry. *Methods Mol. Biol.* **2018**, *1764*, 173–183. [\[CrossRef\]](#)
125. Jones, A.X.; Cao, Y.; Tang, Y.L.; Wang, J.H.; Ding, Y.H.; Tan, H.; Chen, Z.L.; Fang, R.Q.; Yin, J.; Chen, R.C.; et al. Improving mass spectrometry analysis of protein structures with arginine-selective chemical cross-linkers. *Nat. Commun.* **2019**, *10*, 3911. [\[CrossRef\]](#)
126. Politis, A.; Stengel, F.; Hall, Z.; Hernández, H.; Leitner, A.; Walzthoeni, T.; Robinson, C.V.; Aebersold, R. A mass spectrometry-based hybrid method for structural modeling of protein complexes. *Nat. Methods* **2014**, *11*, 403–406. [\[CrossRef\]](#)
127. Rey, M.; Dupré, M.; Lopez-Neira, I.; Duchateau, M.; Chamot-Rooke, J. eXL-MS: An enhanced cross-linking mass spectrometry workflow to study protein complexes. *Anal. Chem.* **2018**, *90*, 10707–10714. [\[CrossRef\]](#) [\[PubMed\]](#)
128. Fritzsche, R.; Ihling, C.H.; Götz, M.; Sinz, A. Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 653–658. [\[CrossRef\]](#)
129. Tinnefeld, V.; Venne, A.S.; Sickmann, A.; Zahedi, R.P. Enrichment of cross-linked peptides using charge-based fractional diagonal chromatography (ChaFRADIC). *J. Proteome Res.* **2017**, *16*, 459–469. [\[CrossRef\]](#)
130. Liu, F.; Lössl, P.; Rabbitts, B.M.; Balaban, R.S.; Heck, A.J.R. The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for coexisting respiratory supercomplexes. *Mol. Cell Proteom.* **2018**, *17*, 216–232. [\[CrossRef\]](#)
131. Tan, D.; Li, Q.; Zhang, M.J.; Liu, C.; Ma, C.; Zhang, P.; Ding, Y.H.; Fan, S.B.; Tao, L.; Yang, B.; et al. Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states. *eLife* **2016**, *5*, e12509. [\[CrossRef\]](#)
132. Rampler, E.; Stranzl, T.; Orban-Nemeth, Z.; Hollenstein, D.M.; Hudecz, O.; Schlögelhofer, P.; Mechtler, K. Comprehensive cross-linking mass spectrometry reveals parallel orientation and flexible conformations of plant HOP2-MND1. *J. Proteome Res.* **2015**, *14*, 5048–5062. [\[CrossRef\]](#) [\[PubMed\]](#)
133. Leitner, A.; Reischl, R.; Walzthoeni, T.; Herzog, F.; Bohn, S.; Förster, F.; Aebersold, R. Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol. Cell. Proteom.* **2012**, *11*, M111.014126. [\[CrossRef\]](#) [\[PubMed\]](#)
134. Herzog, F.; Kahraman, A.; Boehringer, D.; Mak, R.; Bracher, A.; Walzthoeni, T.; Leitner, A.; Beck, M.; Hartl, F.U.; Ban, N.; et al. Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* **2012**, *337*, 1348–1352. [\[CrossRef\]](#)
135. Banerjee, S.; Mazumdar, S. Electrospray ionization mass spectrometry: A technique to access the information beyond the molecular weight of the analyte. *Int. J. Anal. Chem.* **2012**, 282574. [\[CrossRef\]](#) [\[PubMed\]](#)
136. Hermans, J.; Ongay, S.; Markov, V.; Bischoff, R. Physicochemical Parameters Affecting the Electrospray Ionization Efficiency of Amino Acids after Acylation. *Anal. Chem.* **2017**, *89*, 9159–9166. [\[CrossRef\]](#) [\[PubMed\]](#)

137. Johnson, D.; Boyes, B.; Fields, T.; Kopkin, R.; Orlando, R. Optimization of data-dependent acquisition parameters for coupling high-speed separations with LC-MS/MS for protein identifications. *J. Biomol. Tech.* **2013**, *24*, 62–72. [[CrossRef](#)] [[PubMed](#)]
138. Xu, C.-M.; Zhang, J.-Y.; Liu, H.; Sun, H.-C.; Zhu, Y.-P.; Xie, H.-W. Advance of peptide detectability prediction on mass spectrometry platform in proteomics. *Chin. J. Anal. Chem.* **2010**, *38*, 286–292. [[CrossRef](#)]
139. Siggers, T.; Gordán, R. Protein-DNA binding: Complexities and multi-protein codes. *Nucleic Acids Res.* **2014**, *42*, 2099–2111. [[CrossRef](#)]
140. Stützer, A.; Welp, L.M.; Raabe, M.; Sachsenberg, T.; Kappert, C.; Wulf, A.; Lau, A.M.; David, S.S.; Chernev, A.; Kramer, K.; et al. Analysis of protein-DNA interactions in chromatin by UV induced cross-linking and mass spectrometry. *Nat. Commun.* **2020**, *11*, 5250. [[CrossRef](#)]
141. Kramer, K.; Sachsenberg, T.; Beckmann, B.M.; Qamar, S.; Boon, K.L.; Hentze, M.W.; Kohlbacher, O.; Urlaub, H. Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. *Nat. Methods* **2014**, *11*, 1064–1070. [[CrossRef](#)]
142. Trendel, J.; Schwarzl, T.; Horos, R.; Prakash, A.; Bateman, A.; Hentze, M.W.; Krijgsvel, J. The human rna-binding proteome and its dynamics during translational arrest. *Cell* **2019**, *176*, 391–403.e19. [[CrossRef](#)]
143. Ide, H.; Shoukamy, M.I.; Nakano, T.; Miyamoto-Matsubara, M.; Salem, A.M. Repair and biochemical effects of DNA-protein crosslinks. *Mutat. Res.* **2011**, *711*, 113–122. [[CrossRef](#)] [[PubMed](#)]
144. Angelov, D.; Stefanovsky, V.Y.; Dimitrov, S.I.; Russanova, V.R.; Keskinova, E.; Pashev, I.G. Protein-DNA crosslinking in reconstituted nucleohistone, nuclei and whole cells by picosecond UV laser irradiation. *Nucleic Acids Res.* **1988**, *16*, 4525–4538. [[CrossRef](#)] [[PubMed](#)]
145. Stützer, A.; Liokatis, S.; Kiesel, A.; Schwarzer, D.; Sprangers, R.; Söding, J.; Selenko, P.; Fischle, W. Modulations of DNA contacts by linker histones and post-translational modifications determine the mobility and modifiability of nucleosomal H3 tails. *Mol. Cell* **2016**, *61*, 247–259. [[CrossRef](#)] [[PubMed](#)]
146. Sharma, K.; Hrle, A.; Kramer, K.; Sachsenberg, T.; Staals, R.H.; Randau, L.; Marchfelder, A.; Van der Oost, J.; Kohlbacher, O.; Conti, E.; et al. Analysis of protein-RNA interactions in CRISPR proteins and effector complexes by UV-induced cross-linking and mass spectrometry. *Methods* **2015**, *89*, 138–148. [[CrossRef](#)] [[PubMed](#)]
147. Veit, J.; Sachsenberg, T.; Chernev, A.; Aicheler, F.; Urlaub, H.; Kohlbacher, O. LFQProfiler and RNP(xl): Open-source tools for label-free quantification and protein-rna cross-linking integrated into proteome discoverer. *J. Proteome Res.* **2016**, *15*, 3441–3448. [[CrossRef](#)] [[PubMed](#)]
148. Kustatscher, G.; Wills, K.L.; Furlan, C.; Rappsilber, J. Chromatin enrichment for proteomics. *Nat. Protoc.* **2014**, *9*, 2090–2099. [[CrossRef](#)]
149. Ramanathan, M.; Porter, D.F.; Khavari, P.A. Methods to study RNA-protein interactions. *Nat. Methods* **2019**, *16*, 225–234. [[CrossRef](#)]
150. Moore, M.J. From birth to death: The complex lives of eukaryotic mRNAs. *Science* **2005**, *309*, 1514–1518. [[CrossRef](#)]
151. Sutherland, B.W.; Toews, J.; Kast, J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J. Mass Spectrom.* **2008**, *43*, 699–715. [[CrossRef](#)]
152. Chu, C.; Zhang, Q.C.; Da Rocha, S.T.; Flynn, R.A.; Bharadwaj, M.; Calabrese, J.M.; Magnuson, T.; Heard, E.; Chang, H.Y. Systematic discovery of Xist RNA binding proteins. *Cell* **2015**, *161*, 404–416. [[CrossRef](#)] [[PubMed](#)]
153. Simon, M.D.; Wang, C.I.; Kharchenko, P.V.; West, J.A.; Chapman, B.A.; Alekseyenko, A.A.; Borowsky, M.L.; Kuroda, M.I.; Kingston, R.E. The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20497–20502. [[CrossRef](#)] [[PubMed](#)]
154. Li, X.; Song, J.; Yi, C. Genome-wide mapping of cellular protein-RNA interactions enabled by chemical crosslinking. *Genom. Proteom. Bioinform.* **2014**, *12*, 72–78. [[CrossRef](#)] [[PubMed](#)]
155. McHugh, C.A.; Guttman, M. RAP-MS: A method to identify proteins that interact directly with a specific rna molecule in cells. *Methods Mol. Biol.* **2018**, *1649*, 473–488. [[CrossRef](#)]
156. Zeng, F.; Peritz, T.; Kannanayakal, T.J.; Kilk, K.; Eiríksdóttir, E.; Langel, U.; Eberwine, J. A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nat. Protoc.* **2006**, *1*, 920–927. [[CrossRef](#)]
157. Tsai, B.P.; Wang, X.; Huang, L.; Waterman, M.L. Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach. *Mol. Cell. Proteom.* **2011**, *10*, M110.007385. [[CrossRef](#)]
158. Matia-González, A.M.; Iadevaia, V.; Gerber, A.P. A versatile tandem RNA isolation procedure to capture in vivo formed mRNA-protein complexes. *Methods* **2017**, *118–119*, 93–100. [[CrossRef](#)] [[PubMed](#)]
159. Reim, A.; Ackermann, R.; Font-Mateu, J.; Kammel, R.; Beato, M.; Nolte, S.; Mann, M.; Russmann, C.; Wierer, M. Atomic-resolution mapping of transcription factor-DNA interactions by femtosecond laser crosslinking and mass spectrometry. *Nat. Commun.* **2020**, *11*, 3019. [[CrossRef](#)]
160. Urdaneta, E.C.; Vieira-Vieira, C.H.; Hick, T.; Wessels, H.H.; Figini, D.; Moschall, R.; Medenbach, J.; Ohler, U.; Granneman, S.; Selbach, M.; et al. Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nat. Commun.* **2019**, *10*, 990. [[CrossRef](#)] [[PubMed](#)]
161. Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B. Quantitative mass spectrometry in proteomics: A critical review. *Anal. Bioanal. Chem.* **2007**, *389*, 1017–1031. [[CrossRef](#)] [[PubMed](#)]

162. Butter, F.; Scheibe, M.; Mörl, M.; Mann, M. Unbiased RNA-protein interaction screen by quantitative proteomics. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10626–10631. [[CrossRef](#)] [[PubMed](#)]
163. Smith, K.C.; Aplin, R.T. A mixed photoproduct of uracil and cysteine (5-S-cysteine-6-hydrouracil). A possible model for the in vivo cross-linking of deoxyribonucleic acid and protein by ultraviolet light. *Biochemistry* **1966**, *5*, 2125–2130. [[CrossRef](#)]
164. Ule, J.; Jensen, K.B.; Ruggiu, M.; Mele, A.; Ule, A.; Darnell, R.B. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **2003**, *302*, 1212–1215. [[CrossRef](#)] [[PubMed](#)]
165. Licatalosi, D.D.; Mele, A.; Fak, J.J.; Ule, J.; Kayikci, M.; Chi, S.W.; Clark, T.A.; Schweitzer, A.C.; Blume, J.E.; Wang, X.; et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **2008**, *456*, 464–469. [[CrossRef](#)]
166. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr.; Jungkamp, A.C.; Munschauer, M.; et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **2010**, *141*, 129–141. [[CrossRef](#)]
167. Kim, B.; Kim, V.N. fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA. *Methods* **2019**, *152*, 3–11. [[CrossRef](#)] [[PubMed](#)]
168. Liu, Z.R.; Wilkie, A.M.; Clemens, M.J.; Smith, C.W. Detection of double-stranded RNA-protein interactions by methylene blue-mediated photo-crosslinking. *RNA* **1996**, *2*, 611–621.
169. Zaman, U.; Richter, F.M.; Hofele, R.; Kramer, K.; Sachsenberg, T.; Kohlbacher, O.; Lenz, C.; Urlaub, H. Dithiothreitol (DTT) acts as a specific, UV-inducible cross-linker in elucidation of protein-rna interactions. *Mol. Cell. Proteom.* **2015**, *14*, 3196–3210. [[CrossRef](#)]
170. Wower, I.; Wower, J.; Meinke, M.; Brimacombe, R. The use of 2-iminothiolane as an RNA-protein cross-linking agent in Escherichia coli ribosomes, and the localisation on 23S RNA of sites cross-linked to proteins L4, L6, L21, L23, L27 and L29. *Nucleic Acids Res.* **1981**, *9*, 4285–4302. [[CrossRef](#)]
171. Bäumert, H.G.; Sköld, S.E.; Kurland, C.G. RNA-protein neighbourhoods of the ribosome obtained by crosslinking. *Eur. J. Biochem.* **1978**, *89*, 353–359. [[CrossRef](#)] [[PubMed](#)]
172. Noller, H.F.; Chaires, J.B. Functional modification of 16S ribosomal RNA by kethoxal. *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 3115–3118. [[CrossRef](#)]
173. Zhang, Q.; Crosland, E.; Fabris, D. Nested Arg-specific bifunctional crosslinkers for MS-based structural analysis of proteins and protein assemblies. *Anal. Chim. Acta* **2008**, *627*, 117–128. [[CrossRef](#)] [[PubMed](#)]
174. Zhang, Q.; Yu, E.T.; Kellersberger, K.A.; Crosland, E.; Fabris, D. Toward building a database of bifunctional probes for the MS3D investigation of nucleic acids structures. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1570–1581. [[CrossRef](#)] [[PubMed](#)]
175. Karplus, M.; McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652. [[CrossRef](#)] [[PubMed](#)]
176. Hollingsworth, S.A.; Dror, R.O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *19*, 1129–1143. [[CrossRef](#)]
177. Padariya, M.; Kalathiya, U.; Houston, D.R.; Alfaro, J.A. Recognition Dynamics of Cancer Mutations on the ERp57-Tapasin Interface. *Cancers* **2020**, *12*, 737. [[CrossRef](#)]
178. Kalathiya, U.; Padariya, M.; Pawlicka, K.; Verma, C.S.; Houston, D.; Hupp, T.R.; Alfaro, J.A. Insights into the Effects of Cancer Associated Mutations at the UPF2 and ATP-Binding Sites of NMD Master Regulator: UPF1. *Int. J. Mol. Sci.* **2019**, *20*, 5644. [[CrossRef](#)]
179. Kapla, J.; Engström, O.; Svensson, B.; Wohler, J.; Widmalm, G.; Maliniak, A. Molecular dynamics simulations and NMR spectroscopy studies of trehalose-lipid bilayer systems. *Phys. Chem. Chem. Phys.* **2015**, *17*, 22438–22447. [[CrossRef](#)] [[PubMed](#)]
180. Zhang, L.; Bouguet-Bonnet, S.; Buck, M. Combining NMR and molecular dynamics studies for insights into the allostery of small GTPase-protein interactions. *Methods Mol. Biol.* **2012**, *796*, 235–259. [[CrossRef](#)] [[PubMed](#)]
181. Brodie, N.I.; Popov, K.I.; Petrotchenko, E.V.; Dokholyan, N.V.; Borchers, C.H. Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. *Sci. Adv.* **2017**, *7*, e1700479. [[CrossRef](#)] [[PubMed](#)]
182. Shirvanyants, D.; Ding, F.; Tsao, D.; Ramachandran, S.; Dokholyan, N.V. Discrete molecular dynamics: An efficient and versatile simulation method for fine protein characterization. *J. Phys. Chem. B* **2012**, *116*, 8375–8382. [[CrossRef](#)] [[PubMed](#)]
183. Bonomi, M.; Heller, G.T.; Camilloni, C.; Vendruscolo, M. Principles of protein structural ensemble determination. *Curr. Opin. Struct. Biol.* **2017**, *42*, 106–116. [[CrossRef](#)] [[PubMed](#)]
184. Belsom, A.; Schneider, M.; Fischer, L.; Brock, O.; Rappsilber, J. Serum albumin domain structures in human blood serum by mass spectrometry and computational biology. *Mol. Cell. Proteom.* **2016**, *15*, 1105–1116. [[CrossRef](#)] [[PubMed](#)]
185. Boomsma, W.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K. Combining experiments and simulations using the maximum entropy principle. *PLoS Comput. Biol.* **2014**, *10*, e1003406. [[CrossRef](#)]