Anal. Chem. **2010,** *82,* 2636–2642

Chemical Cross-Linking and Mass Spectrometry As a Low-Resolution Protein Structure Determination Technique

Pragya Singh, Alexandre Panchaud, and David R. Goodlett*

Department of Medicinal Chemistry, University of Washington, Seattle, Washington 98195-7610

Protein complexes are the foundation of a majority of cellular processes. Although a large number of protein complexes have been identified through biochemical experiments, the precise molecular details and three-dimensional structures are available for only a small fraction. Chemical cross-linking coupled with mass spectrometry (CXMS) has gained popularity in recent years for characterization of inter- and intraprotein interactions in protein complexes. This perspective provides a comprehensive and critical overview of CXMS strategies employed for structural elucidation of protein complexes. We evaluate the challenges associated with CXMS techniques with special emphasis on data analysis. As sensitivity, mass resolution, mass accuracy and ease of use of mass spectrometers have improved, the complexity of processing and interpreting CXMS data has become the central problem to be addressed. We review here a number of computer programs available to address these problems.

In recent years numerous efforts have been made to advance the role of mass spectrometry in analyzing protein complexes. While most analyses are still carried out on purified complexes in vitro, the potential of mass spectrometry to "visualize" protein interactions in vivo remains a compelling goal. A better understanding of protein complex configurations is important because all fundamental biological processes are controlled by sets of proteins that interact with each other to form transient or stable protein complexes. The regulatory impact of protein−protein interactions on cellular functions accords them prime importance in biological research. This is in part due to dramatic protein conformational changes that occur when proteins interact. Thus, defining protein−protein interactions is key to characterizing protein complexes.

Significant progress has been made toward determining three-dimensional structures of individual proteins by traditional methods like X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR). While these methods provide very high resolution on purified proteins, they have limitations when analyzing multiprotein complexes. Although MS methods cannot compete at the level of detail provided by the high-resolution and global information offered by NMR and X-ray crystallography, they are more tolerant in terms of sample concentration and purity. When coupled with chemical cross-linking, mass spectrometry (i.e., CXMS) can provide site-specific, low-resolution distance constraints with sample quantities that are 2−3 orders of magnitude less than that required for NMR or X-ray crystallography and with faster experimental turnaround times. Additionally, MS analysis can be performed in vitro under conditions that mimic native protein environments, which is not possible with the high-resolution techniques. CXMS methods are also now beginning to push into the realm of in vivo analysis.[1,2] The ability to conduct CXMS experiments in vivo is critical to the future of the field because what we really seek to understand are the real time compositions and spatial distributions of proteins under various biologically relevant conditions.

CXMS analysis of a protein or a protein complex involves formation of covalent bonds between proteins so that the interaction between two amino acid residues close in space is preserved. After chemical conjugation of the two proteins, the cross-linked complex (Figure 1A.1) or the digested peptides (Figure 1A.2) are analyzed by various MS methods. While such MS methods have become routine for characterizing unmodified and post-translationally modified proteins, their application to cross-linked proteins remains limited due in part to a lack of standardized preparatory and informatics methods. To begin with, selection of the most appropriate chemical cross-linking reagent and optimization of cross-linking conditions for even a small complex can be tedious. However, the most challenging aspect of CXMS analysis is the large assortment of peptides resulting from proteolytic digestion of cross-linked proteins (Figure 1B) because the abundance of cross-linked peptides in this mixture is typically very low relative to the linear peptides. This alone makes detection of these nonstoichiometric, cross-linked peptide components extremely difficult. Two areas where progress in detecting these species has been made are design of the cross-linking reagents and bioinformatics that facilitate detection of cross-linked peptides. These two topics are the main focus of this perspective that aims to provide readers an analytical overview of the experimental techniques and bioinformatics strategies existing in the field of CXMS.

* To whom correspondence should be addressed. Dr. David R. Goodlett, University of Washington, Department of Medicinal Chemistry, Box 357610, Seattle, WA 98195-7610. Phone: 206-616-4586. Fax: 206-685-3252. E-mail: goodlett@u.washington.edu.

(1) Suchanek, M.; Radzikowska, A.; Thiele, C. *Nat. Meth.* **2005**, *2*, 261–268.
(2) Zhang, H.; Tang, X.; Munske, G. R.; Tolic, N.; Anderson, G. A.; Bruce, J. E. *Mol. Cell. Proteomics* **2009**, *8*, 409–420.

10.1021/ac1000724
Published on Web 03/08/2010

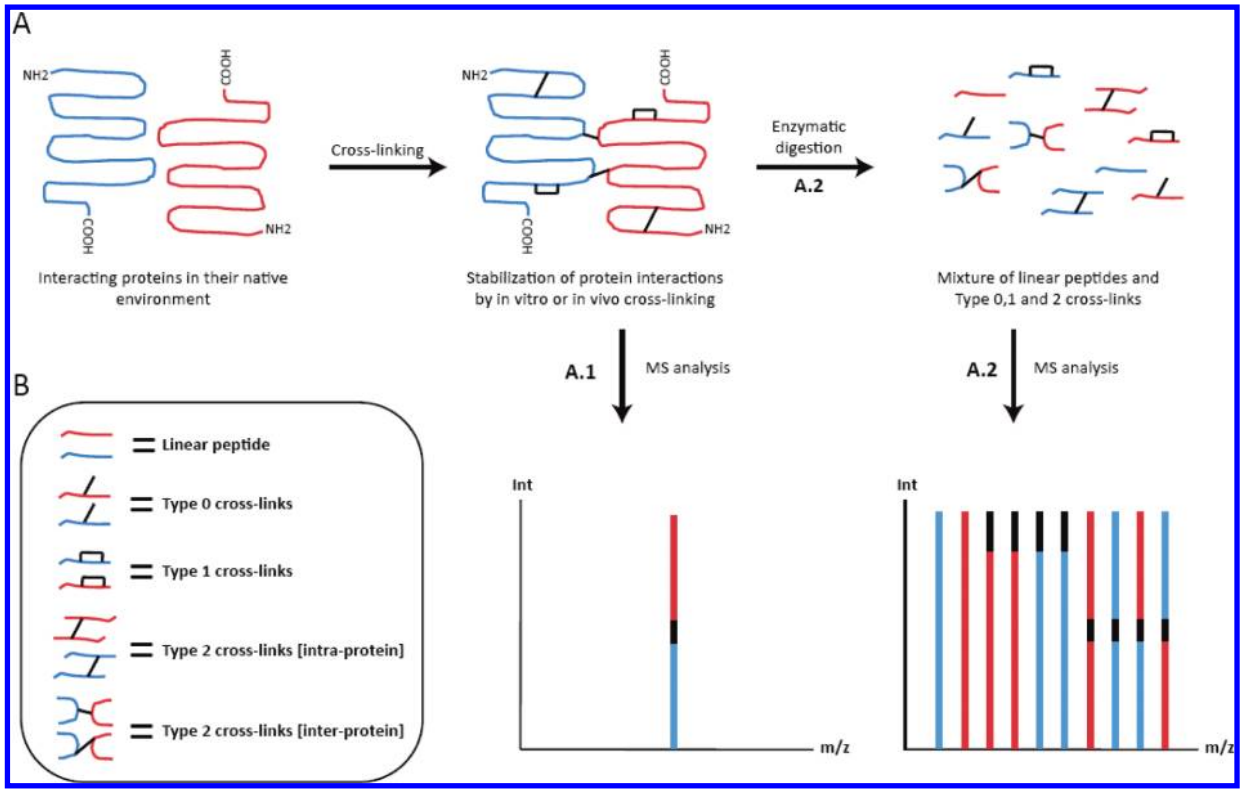

Figure 1. General scheme showing CXMS analysis for characterizing protein−protein interactions: (A.1) top-down approach, (A.2) bottom-up approach. (B) Different products generated from proteolytic digestion of cross-linked complexes. A detailed discussion on nomenclature of cross-linked peptides, as presented in this figure, is available in the Supporting Information.

STRATEGIES USED FOR DETECTION OF CROSS-LINKED PEPTIDES

A large number of chemical cross-linkers with varying chemical specificities and spacer-arm lengths are now available.[3−6] Detailed descriptions of these cross-linkers along with their advantages and limitations are provided in the Supporting Information (Figure S1). Historically, the primary reasons for cross-linking protein complexes were to (i) facilitate analysis of transient complexes by stabilizing weak interactions and (ii) estimate proximity of interacting residues. However, it was soon realized that reagent design could be manipulated to facilitate detection of cross-linked peptides. Given the wide array of reaction products generated by cross-linking experiments, reagents that aid detection of cross-linked peptides, a problem akin to finding the proverbial needle in a haystack, was enthusiastically greeted in the community. Cross-linking reagents have been designed with specific functionalities within the spacer arm that allow selective monitoring of cross-linked peptides, such as stable isotopes, radioactive isotopes, and fluorophores. Affinity tags have also been incorporated to allow both purification and identification. Following are descriptions of some of these sophisticated detection strategies.

Isotope Labeling. Incorporation of stable isotopes in peptides and proteins can be used to distinguish cross-linked peptides from linear peptides. Isotopic derivatization of proteolyzed peptides with heavy and light reagents[7] and expression of target protein in heavy and light (^{14}N-labeled) versions[8] have been used successfully to identify cross-linked peptides via MS. Another isotope-labeling strategy involves tryptic digestion of cross-linked peptides in the presence of ^{18}O-labeled water.[9] These strategies are based on the fact that cross-linked peptides, by virtue of having two N- and C-termini, produce a characteristic isotopic pattern in MS analysis that can be distinguished from those produced by noncross-linked peptides or dead-end products. However, these strategies involve labor intensive multistep chemistries, incomplete labeling, or back-exchange of the label and excessive sample handling. Despite these drawbacks, the advantage of isotopic-labeling is that the method is independent of the cross-linking reagent used, i.e., it can be used with any cross-linker and can also be applied to characterize native cross-links, such as disulfide bonds.

The first study to use stable isotope labeled cross-linkers was reported by Muller et al. They used deuterated and undeuterated versions of three different reagents to identify interacting sites in a protein complex.[10] This study was soon followed by numerous studies using d0/d4 labeled cross-linkers and computer programs to handle data generated from such experiments (see Bioinformatics: From Manual to Automated Interpretation). Given the

(3) Back, J. W.; de Jong, L.; Muijsers, A. O.; de Koster, C. G. *J. Mol. Biol.* **2003**, *331*, 303–313.
(4) Sinz, A. *J. Mass Spectrom.* **2003**, *38*, 1225–1237.
(5) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25*, 663–682.
(6) Pierce. Rockford, IL, 2003.
(7) Chen, X.; Chen, Y. H.; Anderson, V. E. *Anal. Biochem.* **1999**, *273*, 192–203.
(8) Taverner, T.; Hall, N. E.; O'Hair, R. A. J.; Simpson, R. J. *J. Biol. Chem.* **2002**, *277*, 46487–46492.
(9) Back, J. W.; Notenboom, V.; de Koning, L. J.; Muijsers, A. O.; Sixma, T. K.; de Koster, C. G.; de Jong, L. *Anal. Chem.* **2002**, *74*, 4417–4422.
(10) Muller, D. R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M. O. *Anal. Chem.* **2001**, *73*, 1927–1934.

popularity of deuterated reagents, many of these are now available commercially.[6] In addition to deuterium, ^{18}O has also been incorporated into cross-linking reagents to create isotopic patterns unique to cross-linked peptides.[6,11] One benefit of using ^{18}O as the isotope label is that it does not produce a chromatographic-isotope effect like deuterium during LC−MS/MS analysis.

Chemically Labile Reagents. A chemically labile moiety within the reagent can be used for identifying cross-linked peptides via MS. The most commonly used moiety for this purpose is a disulfide bond (RS−SR) (e.g., 3,3′-dithiobis(sulfosuccinimidyl-propionate) (DTSSP), dithiobis(succinimidyl propionate) (DSP), succinimidyl 6-(3′-[2-pyridyl-dithio]-propionamido)hexanoate) (LC-SPDP)).[6] The cross-linked peptides generated using such chemistry, when reduced in the presence of a thiol reagent, produce two corresponding cross-linker-modified peptides. A differential analysis of precursor ion spectra from nonreduced and reduced mixtures is then used to identify the cross-linked candidates. However, presence of cysteines in proteins can complicate this method due to disulfide exchange with the reagent. This problem can be overcome by using reagents that incorporate other chemically labile moieties besides disulfide bonds. For example, disuccinimidyl tartarate (DST) and 1,4-bis-maleimidyl-2,3-dihydroxybutane (BMDB) (Pierce, Rockford, IL) both of which can be cleaved by sodium meta-periodate.[6]

MS Labile Reagents. Back et al. designed a MS labile cross-linker, *N*-benzyliminodiacetoyloxysuccinimid (BID), that yields a benzyl cation ($m/z = 91$) as a marker ion during low-energy collisionally-induced dissociation (CID).[12] Since the benzyl group is incorporated as a side chain of the reagent, dissociation leaves the cross-link between the two peptides intact. So, while this marker ion allows the selection of potential cross-linked candidates, it does not facilitate sequence identification of the two peptides that are cross-linked. Another set of MS labile reagents fragments to produce two individual peptides from the cross-linked species. This set includes CID-CXL reagents designed by Solderblom et al.[13,14] These amine reactive reagents contain a single MS labile Aspartyl−Prolyl bond within the linker which readily dissociates under low-energy CID conditions to generate two peptides that can be subsequently fragmented to provide sequence information. Similarly, an ionic reagent has been reported by Lu et al.,[15] which has a labile C−S bond that fragments during CID. The strength of CID-CXL reagents and the ionic reagent lies in the fact that any conventional database search engine can be used to analyze the data. However, these reagents are not available commercially and need to be synthesized in-house. Tang and coworkers developed a series of MS-labile reagents termed protein interaction reporters (PIRs).[16,17] PIRs contain two labile bonds within the spacer arm that can dissociate to release a marker ion as well as modified versions of the two peptides that were originally cross-linked. The released peptides can be subsequently sequenced by another stage of tandem MS. Though commercially unavailable, these reagents provide a sophisticated strategy to detect potential cross-linked candidates and determine their sequence in a single analysis. PIRs have also been coupled with other chemical features, such as affinity tagging, photocleavable groups, to produce a series of trifunctional reagents.[18] Recently, a "chromogenic" reagent that allows selective dissociation of cross-linked peptides during infrared multiphoton dissociation (IRMPD) was reported.[19] Since this strategy involves tedious comparison of reconstructed total ion chromatograms before and after IR exposure for cross-link identification, it appears to be best suited to purified mixtures containing a small number of proteins.

PURIFICATION OF CROSS-LINKED PEPTIDES: FROM COMPLEX TO SIMPLIFIED MIXTURES

Regardless of the cross-linker used, each cross-linking strategy eventually faces the daunting task of purifying cross-linked peptides from the large majority of linear peptides prior to MS analysis. Some of the purification techniques used for this purpose are illustrated in Figure 2. For simplified mixtures with relatively few proteins, it is often useful to enrich the cross-linked products at the protein level by SDS-PAGE and/or size-exclusion chromatography (SEC). Both techniques can be followed by in-gel or in-solution proteolysis, respectively, to identify the proteins and cross-link sites by MS. At the peptide level, peptide physio-chemical properties, e.g., isoelectric point or peptide length/mass, can be used to enrich for cross-linked peptides. For example, it has been shown that a majority of cross-linked peptides have higher molecular weights (Figure 3A) than linear peptides and that they often possess charge-states ≥4+.[20,21] Thus, cross-linked peptides can be enriched by strong cation exchange chromatography (SCX) under acidic conditions. Recently, we found that high charge-state peptides can also be enriched using C4 instead of C18 reversed phase chromatography; a 60% increase in the number of spectra acquired from high charge-state (≥4+) precursors using C4 chromatography compared to C18 was observed. Finally, we and others have exploited the high charge-state of cross-linked peptides by targeted tandem MS acquisition of ions with charge-states ≥4+[20,21] (Figure 3B).

BIOINFORMATICS: FROM MANUAL TO AUTOMATED INTERPRETATION

Even with advances in mass spectrometric instrumentation and improvements in cross-linking reagents and separation techniques, the complexity and volumes of data generated from CXMS experiments remain formidable. The problem of extracting structural information from the data has been met largely by

(11) Collins, C. J.; Schilling, B.; Young, M.; Dollinger, G.; Guy, R. K. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4023–4026.
(12) Back, J. W.; Hartog, A. F.; Dekker, H. L.; Muijsers, A. O.; de Koning, L. J.; de Jong, L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 222–227.
(13) Soderblom, E. J.; Bobay, B. J.; Cavanagh, J.; Goshe, M. B. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3395–3408.
(14) Soderblom, E. J.; Goshe, M. B. *Anal. Chem.* **2006**, *78*, 8059–8068.
(15) Lu, Y.; Tanasova, M.; Borhan, B.; Reid, G. E. *Anal. Chem.* **2008**, *80*, 9279–9287.
(16) Tang, X.; Munske, G. R.; Siems, W. F.; Bruce, J. E. *Anal. Chem.* **2005**, *77*, 311–318.
(17) Zhang, H.; Tang, X.; Munske, G. R.; Zakharova, N.; Yang, L.; Zheng, C.; Wolff, M. A.; Tolic, N.; Anderson, G. A.; Shi, L.; Marshall, M. J.; Fredrickson, J. K.; Bruce, J. E. *J. Proteome Res.* **2008**, *7*, 1712–1720.
(18) Chowdhury, S. M.; Munske, G. R.; Tang, X.; Bruce, J. E. *Anal. Chem.* **2006**, *78*, 8183–8193.
(19) Gardner, M. W.; Vasicek, L. A.; Shabbir, S.; Anslyn, E. V.; Brodbelt, J. S. *Anal. Chem.* **2008**, *80*, 4807–4819.
(20) Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller, L.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R. *Nat. Meth.* **2008**, *5*, 315–318.
(21) Singh, P.; Shaffer, S. A.; Scherl, A.; Holman, C.; Pfuetzner, R. A.; Larson Freeman, T. J.; Miller, S. I.; Hernandez, P.; Appel, R. D.; Goodlett, D. R. *Anal. Chem.* **2008**, *80*, 8799–8806.

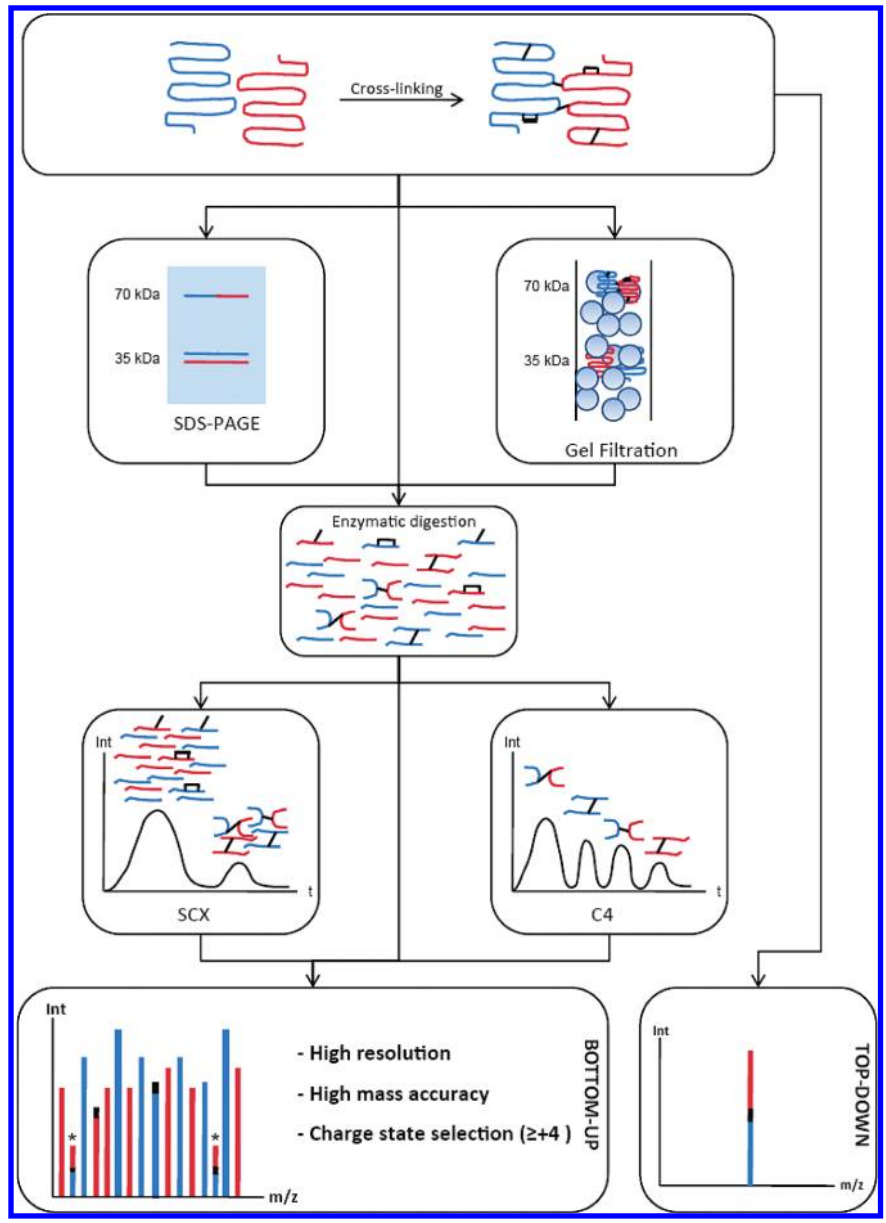

Figure 2. Analytical strategies used for separation of cross-linked products before and after proteolytic digestion.

laboratory-specific software developments; see the Supporting Information for a summary table of all available tools. Several groups have developed proprietary (nonpublic) algorithms for this purpose. Other programs are commercially available, and others are still only available as compiled code. The first generation of software is based only on precursor ion data (Table S1.A in the Supporting Information), which work by creating a list of theoretical cross-linked peptides from protein(s) of interest based on user specified proteases and cross-linking reagents. Experimental m/z values are matched, within a defined mass tolerance, to theoretical values to identify potential cross-linked candidates. While the precursor-ion-based methods work, they are time-consuming, are computationally expensive, and have high false positive rates due to the exponential possibilities from peptide−peptide combinations. Moreover, the results from these programs always require validation based on tandem MS data.

The next generation of cross-linking software consisted of programs that have dual functionality using precursor ion data to predict cross-linked peptide candidates followed by tandem MS data to validate those candidates (Table S1.B in the Supporting Information). Two such tools are Pro-Cross-link[22] and XLINK,[23] both of which require two separate liquid chromatography−mass spectrometry (LC−MS) and and liquid chromatography−tandem mass spectrometry (LC−MS/MS) data sets. These programs make use of isotope labeling information from precursor ion data to select the cross-linked candidates. The selected candidates are investigated and confirmed based on precursor ion spectra and tandem mass spectra, and the amino acid sequences are assigned for the cross-linked peptide pairs. Pro-Cross-link was originally developed for LC−electrospray (ESI)-MS data while XLINK was developed for LC−matrix-assisted laser desorption ionization (MALDI)-MS data. While MALDI analysis allows one to use the

(22) Gao, Q.; Xue, S.; Doneanu, C. E.; Shaffer, S. A.; Goodlett, D. R.; Nelson, S. D. *Anal. Chem.* **2006**, *78*, 2145–2149.

(23) Seebacher, J.; Mallick, P.; Zhang, N.; Eddes, J. S.; Aebersold, R.; Gelb, M. H. *J. Proteome Res.* **2006**, *5*, 2270–2282.

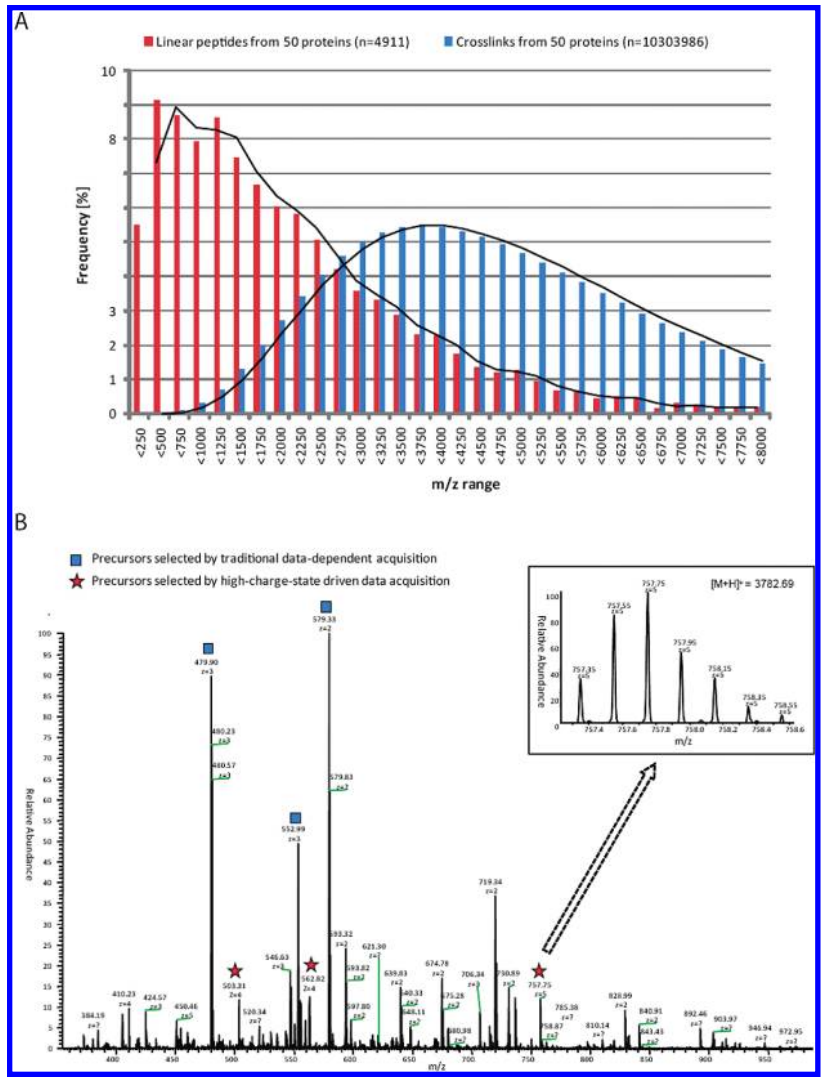

Figure 3. (A) Theoretical weight distribution of cross-linked peptides and linear peptides in 50 human proteins including Cyp2E1 and *b5*. Proteins were cross-linked with EDC and digested with trypsin in silico. (B) Precursor ion spectrum showing selection of low abundance 4+ and 5+ charge-state precursors over abundant 2+ and 3+ charged precursors during the high charge-state driven data acquisition.

same spot on the MALDI plate for prescreening at the precursor ion level as well as the confirmatory analysis via tandem mass spectrometry, the MALDI-based method is not as efficient for fragmenting large cross-linked peptides.

The latest generation of cross-linking software includes programs designed to identify cross-linked peptides directly by searching tandem mass spectra generated in shotgun proteomic experiments (Table S1.C in the Supporting Information). A number of such sophisticated programs have emerged in the last 2 years underscoring the need for automated high-throughput tools for cross-link analysis. One such program, X!Link, which relies on high mass accuracy of precursor ions and good quality tandem mass spectral data to identify cross-linked peptides,[24] does not require isotope labeling for preselection of potential candidates. For each tandem mass spectrum, X!Link first queries the precursor mass to find if it matches the mass of a cross-linked peptide pair. If it does, then X!Link generates a theoretical fragment ion list for the cross-linked peptide pair and compares it to the experimental fragments to assign the number of matched peaks and a score based on matched fragments. In summary, this program elegantly incorporates the strategies of precursor-ion-based cross-link identification tools and tandem MS assignment programs in one automated tool. However, it does suffer from some shortcomings. For example, separate Sequest or Mascot searches are required to analyze dead-ends and intrapeptide cross-links and it cannot be used for analyzing protein complexes comprising more than two proteins. Additionally, X!Link only considers lysine reactive cross-linkers, which precludes the use of several other important reagents, such as 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), formaldehyde, etc., and for 4+ and higher charge-state precursors, X!Link only accepts their doubly and triply charged fragment ions which automatically puts these precursors at a disadvantage because they often contain fragments ions with >3+ charge-state.

Another approach to automate the process of cross-link identification involves database searching,[25] a strategy similar to that used by programs like Sequest and Mascot for protein identification in LC−MS/MS experiments. As straightforward as

(24) Lee, Y. J.; Lackner, L. L.; Nunnari, J. M.; Phinney, B. S. *J. Proteome Res.* **2007**, *6*, 3908–3917.

(25) Maiolica, A.; Cittaro, D.; Borsotti, D.; Sennels, L.; Ciferri, C.; Tarricone, C.; Musacchio, A.; Rappsilber, J. *Mol. Cell. Proteomics* **2007**, *6*, 2200–2211.

it sounds, this process is significantly challenging for cross-linked peptides. The biggest challenge lies in creating a special cross-link database that can be used by standard database search tools such as Mascot or Sequest. Such a database strategy was presented by Mailoca et al. who used in silico digestion of all theoretical cross-links of interacting proteins to create a concatenated cross-linked protein database. This approach, however, generates a very large sequence search space of which only a fraction represents feasible cross-linked peptide pairs. This increases both search time and false-positive rates and necessitates a customized scoring scheme to sort the data. Still their strategy has two main advantages: (1) once the cross-linked database has been created, any database search engine can perform the database search, and (2) a decoy or a control search can be used for estimating false positive rates.

Recently, a very elegant program xQuest was published for detecting cross-linked peptides from large protein databases.[20] With the use of xQuest, searches can be performed in one of two modes (i) enumeration mode where all possible peptide combinations are stored as an index, providing an exhaustive search that is limited to 100 proteins and (ii) ion-tag mode where each peptide in the database is first associated with specific fragment ion masses as an ion index, followed by enumeration of peptide−peptide combinations consistent with the ion indices. Candidates from both enumeration mode and ion-tag mode are scored using a scoring routine similar to Sequest. xQuest stands apart from the other tools in that it allows comprehensive analysis of cross-linked spectra from complex mixtures. While it works very well, there are notable disadvantages. For example, the algorithm requires isotope labeling to work well for large database searches, and the scoring routine can be computationally expensive, especially for complex mixtures with thousands of spectra. Also, there appears to be some room for improvement in the process of fragment ion searching because the method uses tandem mass spectra acquired in a low-resolution LTQ ion trap which limits confidence in a sequence match.

Our group recently reported yet another cross-link strategy that utilizes data from a single LC−MS/MS run[21] providing significant advantage for samples of limited availability. This method, designed to be as generic as possible, is based on an open-modification search engine Popitam (freely accessible at www.expasy.ch/tools/popitam/). Unlike the programs described above, this strategy does not use precursor mass as a filter for selecting tandem mass spectra for further analysis. Furthermore, it does not require prefractionation or isotopic labeling of cross-linked peptides before analysis. Instead, it relies only on fragmentation of high-charge-state ($\geq [M + 4H]^{4+}$) precursors (Figure 3B) and acquisition of precursors as well as fragment ions at high mass accuracy. Conceptually, the approach considers each peptide in a cross-link to be post-translationally modified with an unknown mass (i.e., the second peptide) at an unknown amino acid. A cross-linked peptide is considered as a match by Popitam[26] if two sequence tags corresponding to two unique peptides are identified from the same tandem mass spectrum and if the sum of those peptide masses and the cross-linker matches the precursor mass (within 5 ppm). This method is very specific because high mass accuracy of precursors and fragment ions as well as the specificity of a derived sequence tag is sufficient to provide a small number of cross-linked candidates from thousands of spectra.[21]

IN VIVO CROSS-LINKING: FROM STABILIZATION TO STRUCTURE

The final goal of efforts mentioned so far is to observe proteins interacting in their native or in vivo environment. Thus, before we conclude, we review recent in vivo cross-linking efforts. The term "in vivo cross-linking" can often be ambiguous in the context of protein−protein interaction studies. There are two main applications of cross-linking which make it interesting in biological studies. The first, which is the primary focus of this perspective, consists of chemically linking residues in close proximity to estimate distance constraints. The second involves stabilizing transient protein interactions using chemical reagents, prior to sample preparation and MS analysis, to identify components of the protein complex. In this case, the goal is only to identify interacting partners in a protein complex rather than their relative juxtaposition. Most of the studies reporting successful in vivo cross-linking refer to the latter application because the former remains challenging.

Although, in theory many cross-linking reagents can be used to stabilize protein interactions, only a few have been successfully applied in vivo to identify interacting partners. Formaldehyde has been used in a refined technique termed time-controlled transcardiac perfusion cross-linking (tcTPC) to study the mouse brain interactome.[27,28] In this report, the cross-linker was pumped through the circulatory system for a short duration to capture the physiological interactions before disrupting the target tissue and purifying the protein complexes. Formaldehyde, because of its small size, readily permeates cell walls and membranes making it very promising for such studies. Use of photoreactive amino acids (L-photoleucine and L-photomethionine) is another novel strategy for studying protein interactions in their native environment. After incorporation in the cells during protein synthesis, cells are UV irradiated to create cross-links.[1] While used only for identifying interacting partners, this strategy holds great promise for large scale interactome studies comprising structural analysis.

Finally, Zhang and colleagues recently published what appears to be the first paper using in vivo chemical cross-linking for structural elucidation.[2] In their study, PIRs were used for intact cell labeling and a two-stage MS strategy was utilized for analysis of contact/binding interfaces in living cells. Their results clearly display the potential of CXMS for in vivo studies. The studies mentioned in this section demonstrate that with continued development and refinement, CXMS will eventually provide the community with a robust tool to probe and elucidate in vivo cellular processes.

(26) Hernandez, P.; Gras, R.; Frey, J.; Appel, R. D. *Proteomics* **2003**, *3*, 870–878.

(27) Schmitt-Ulms, G.; Hansen, K.; Liu, J.; Cowdrey, C.; Yang, J.; DeArmond, S. J.; Cohen, F. E.; Prusiner, S. B.; Baldwin, M. A. *Nat. Biotechnol.* **2004**, *22*, 724–731.

(28) Bai, Y.; Markham, K.; Chen, F.; Weerasekera, R.; Watts, J.; Horne, P.; Wakutani, Y.; Bagshaw, R.; Mathews, P. M.; Fraser, P. E.; Westaway, D.; St. George-Hyslop, P.; Schmitt-Ulms, G. *Mol. Cell. Proteomics* **2008**, *7*, 15–34.

WHERE DOES THE CXMS FIELD STAND RELATIVE TO NEEDS OF THE BIOCHEMICAL COMMUNITY?

The strategies and tools outlined in this perspective hold great promise for tackling challenges presented to the biochemical community to define and characterize protein complexes. The final objective when using these strategies is often to determine biologically relevant protein−protein interactions associated with biological activity of proteins. The data generated from CXMS experiments provide a means to refine knowledge of the associations and distance constraints needed to correlate the dynamics of structure−function relationships. Thus, the many CXMS pipelines reviewed here are often only the beginning of a series of experiments to verify, through complementary techniques, the proposed structures and functions.

To this end, the interplay between molecular modeling and CXMS generated hypotheses is promising. Computer generated theoretical models of a given complex can be screened using the distance constraints provided by CXMS to eliminate inconsistent structures. This can be used iteratively to improve accuracy of the computer generated models. A new software platform called MSX-3D allows one to conduct such associations of in silico predicted protein structures and CXMS data.[29] The MSX-3D platform facilitates testing theoretical models to validate, refute and redefine structures based on CXMS data. To our knowledge, this is the first attempt to bridge the gap between CXMS and protein structural modeling. Refinement of this and future tools with similar objectives will expedite the knowledge transfer from CXMS data to biological structure−function relationships.

ACKNOWLEDGMENT

P.S. and A.P. contributed equally to this work.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review January 11, 2010. Accepted February 23, 2010.

AC1000724

(29) Heymann, M.; Paramelle, D.; Subra, G.; Forest, E.; Martinez, J.; Geourjon, C.; Deleage, G. *Bioinformatics* **2008**, *24*, 2782–2783.