April 24, 2007

Shotgun Cross-Linking Analysis for Studying Quaternary and Tertiary Protein Structures

Young Jin Lee, *University of California, Davis*
Laura L. Lackner, *University of California, Davis*
Jodi M. Nunnarl, *University of California, Davis*
Brett S. Phinney, *University of California, Davis*

research articles Journal of proteome research

Shotgun Cross-Linking Analysis for Studying Quaternary and Tertiary Protein Structures

Young Jin Lee,*,† Laura L. Lackner,‡ Jodi M. Nunnari,‡ and Brett S. Phinney†

Proteomics Core Facility, Genome Center, and Molecular and Cellular Biology, University of California, Davis 95616

Received April 24, 2007

We developed a new approach that employs a novel computer algorithm for the sensitive and high-throughput analysis of tertiary and quaternary interaction sites from chemically cross-linked proteins or multi-protein complexes. First, we directly analyze the digests of the chemically cross-linked proteins using only high-accuracy LC−MS/MS data. We analyze these data using a computer algorithm, we term X!Link, to find cross-links between two peptides. Our algorithm is rapid, taking only a few seconds to analyze ~5000 MS/MS spectra. We applied this algorithm to analyze cross-linked sites generated chemically using the amino specific reagent, BS3, in both cytochrome c and the mitochondrial division dynamin mutant, Dnm1G385D, which exists as a stable homodimer. From cytochrome c, a well-established test protein, we identified a total of 31 cross-links, 21 interpeptide and 10 intrapeptide cross-links, in 257 MS/MS spectra from a single LC−MS/MS data set. The high sensitivity of this technique is indicated by the fact that all 19 lysines in cytochrome c were detected as a cross-link product and 33% of all the Lys pairs within 20 Å were also observed as a cross-link. Analysis of the cross-linked dimeric form of Dnm1G385D identified a total of 46 cross-links, 38 interpeptide and 8 intrapeptide cross-links, in 98 MS/MS spectra in a single LC−MS/MS data set. These results represent the most abundant cross-links identified in a single protein or protein dimer to date. Statistical analysis suggests a 1% false discovery rate after optimization of filtering parameters. Further analysis of the cross-links identified using our approach indicates that careful manual inspection is important for the correct assignment of cross-linking sites when multiple cross-linkable sites or several similar sequences exist. In summary, we have developed a sensitive MS-based approach to identify peptide−peptide cross-links that does not require isotopic labeling or comparison with non-cross-linked controls, making it faster and simpler than current methodologies.

Keywords: shotgun • tertiary protein structure • quaternary protein structure • protein−protein interaction • cross-link • cross-linking reagent • mass spectrometry • proteomics • protein complex

Introduction

Mass spectrometric (MS) analysis of chemically cross-linked proteins has great potential for the structural analysis of protein−protein interactions and protein three-dimensional structures. This technique can identify exact interaction sites in protein complexes with relatively small amounts of material. In addition, distance constraints imposed by the arm length of a chemical cross-linker can be used in the modeling of protein three-dimensional structures and can also provide important supplemental information to X-ray or NMR structural data or can provide a low-resolution approach to structure when X-ray or NMR methods are not possible.

Young and co-workers' analysis of bovine fibroblast growth factor was pioneering work to the development of MS for the analysis of protein cross-links.[1] Specifically, they successfully identified 18 Lys−Lys cross-links from a series of MALDI-TOF and MALDI/PSD spectra of LC fractionated digested peptides and improved threading models with cross-link-derived distance constraint errors. Since then, many groups have utilized this approach for various biological questions. However, in its current form, the use of MS to identify peptide cross-links is laborious and time-consuming because cross-linked peptides are present in a mixture containing a much greater abundance of non-cross-linked peptides. Indeed, this task has often been likened *to finding a needle in a hay stack*.

Cross-linking mass spectrometry studies have been based predominantly on comparing the mass spectra of cross-linked samples with those of unmodified controls. These studies use computer programs designed for peptide mass mapping such as GPMAW or PAWS[2,3] or rely on their own programs that compute and list all the possible cross-linked peptide masses (ASAP,[1] FindLink,[4] X-Link,[5] NIH-XL,[6] CLPM,[7] MS-Bridge[8]). This approach has been shown to work, but it is laborious, insensi-

* To whom correspondence should be addressed. Phone, 530-754-9474; Fax, 530-754-8370; E-mail, yojlee@ucdavis.edu.
† Genome Center.
‡ Molecular and Cellular Biology.

10.1021/pr070234i CCC: $37.00
Published on Web 09/14/2007

tive, and difficult to automate. Reliability of the identified cross-links is very important considering that a combination of two theoretical peptides leads to an exponential number of possible cross-linked peptides. It has become essential, therefore, to acquire accurate mass information using Fourier transform mass spectrometery[3,9] and/or to get structural information in MS/MS.[2−5,8,9]

To improve the ease and sensitivity of detecting cross-linked peptides, isotopic labeling has often been adopted in cross-linking or digestion. However, until recently, no high-throughput tool has been available. Recently, two University of Washington groups independently developed high-throughput cross-linking mass spectrometry methods to automate the analysis based on isotopic labeling and the use of LC−ESI or LC−MALDI: Pro-Crosslink[9] and iXLink/doXLink.[10] However, despite this advance, these methods still have limited sensitivity for detecting isotopically labeled peptide patterns and require two-step procedures to identify candidates and confirm MS/MS spectra.

In this paper, we report an easy-to-use, sensitive, and high-throughput *shotgun* method to identify cross-linked peptides by directly analyzing LC−MS/MS data, without any isotopic labeling or analysis of a non-cross-linked control sample. Our data indicate that accurate parent mass and high quality MS/MS spectra obtained with LTQ-FT are sufficient to identify cross-linked peptides in a mixture with non-cross-linked peptides as in other post-translational modification profiling. Shotgun proteomics tools, such as X!Tandem[11] and SEQUEST,[12] are used to search for single peptides modified by chemical cross-linkers, either intrapeptide cross-links or monolinks. For interpeptide cross-link searches, we have developed a computer algorithm, termed X!Link, which is optimized to yield a low false discovery rate. We have successfully applied this new method to identify intra- and interpeptide chemically formed cross-links in cytochrome c and the mitochondrial division dynamin homodimer, Dnm1G385D.

Experimental

Sample Preparation for Cytochrome c. Cytochrome c (horse heart, Sigma, St. Louis, MO) was purified with a spin concentrator (Microcon, 3 kDa MWCO; Millipore, Bedford, MA) in 20 mM phosphate buffer (pH 7.2; Pierce, Rockford, IL). A 20-fold molar excess of BS3 (Bis[sulfosuccinimidyl] suberate, Pierce) was added to 5 μM cytochrome c and incubated for 2 h at ambient temperature (RT). After a buffer change to 50 mM ammonium bicarbonate using a spin concentrator, it was digested with trypsin (Modified trypsin, sequencing grade; Promega, Madison, WI) at a 1:50 enzyme-to-substrate ratio (wt:wt) and incubated overnight at 37 °C. The reaction was quenched with 5% formic acid, and 1 μg of the total digest was loaded on the LC−MS/MS system.

Dnm1G385D Purification and Cross-Linking. Dnm1G385D was purified via an N-terminal 10×His tag as previously described.[13] The cross-linking reaction was performed in 25 mM Hepes (pH 7.0), 25 mM Pipes (pH 7.0), 400 mM NaCl, and 17% DMSO. A 50-fold molar excess of BS3 was added over Dnm1G385D so that the final concentration of BS3 in the reaction was 350 μM and that of Dnm1G385D was 7 μM. The reaction was incubated at RT for 30 min, after which time Tris-Cl (pH 7.5) was added to 50 mM to quench the reaction. The reaction was incubated for an additional 15 min at RT. The Dnm1G385D dimer and monomer bands were resolved by SDS-PAGE and visualized by staining with Coomassie Blue. Dnm1G385D quantitatively exists only as a dimer in solution under our cross-linking condition and monomer band is from dimers uncross-linked between monomer units and denatured by SDS. Approximately 14 μg of total protein was loaded per lane. The dimer band was cut from the gel, reduced with 10 mM DTT at 56 °C for 30 min, alkylated with 55 mM iodoacetamide at RT for an hour in the dark, and digested with trypsin at a 1:50 enzyme-to-substrate ratio (wt:wt) overnight at 37 °C. Half of the digest was used for LC−MS/MS analysis.

LC−MS/MS Data Acquisition. A hybrid high-resolution mass spectrometer, LTQ-FT (Thermo Electron, San Jose, CA), coupled with a NanoAcquity UPLC system (Waters, Milford, MA) was used to acquire LC−MS/MS data. A NanoAcquity UPLC column (100 μm × 10 cm; BEH C18, 1.7 μm, 100 Å; Waters) was used with a 90 min-long gradient (1−10% buffer B for 5 min, 10−35% buffer B for 65 min, 35−70% buffer B for 5 min, 70% buffer for 1 min, 1% buffer B for 14 min) at a flow rate of 1 μL min^{-1}. Buffer A is 0.1% formic acid and buffer B is 0.1% formic acid in acetonitrile. A Fourier transform survey scan at resolution 25 000 is followed by three most intense ions' FT zoom scans at resolution 50 000 and MS/MS scans by LTQ. Dynamic exclusion was used with the following parameters: repeat count, 2; repeat duration, 15 s; exclusion duration, 45 s. An isolation window of 3 Da was used for precursor ion selection and a normalized collision energy of 35% was used for the fragmentation. Singly charged ions are rejected for MS/MS acquisition.

Data Analysis. MASCOT Distiller (v2.1.0.0) was used to deisotope MS/MS spectra. Default parameters for LTQ-FT were used with the following changes: in MS, regridding points per Da of 400, maximum peak charge of 7; in MS/MS, uncentroiding half width of 0.1 Da, uncentroiding and regridding points per Da of 30, minimum peak count of 10, and precursor m/z tolerance to be grouped of 1.2 Da; in peak selection, minimum SN ratio of 2, expected peak width of 0.02 Da.

SEQUEST (Bioworks 3.3, Thermo) and X!Tandem (version 2006.06.01.2, through GPM-XE 2006.09.15, Beavis Informatics, Canada) were used for the chemical cross-linking analysis in a single peptide: monolink ("dead-end") and intrapeptide link ("loop-link"). Only fully tryptic peptides were included in both searches. In X!Tandem, IPI human database (v3.24) added with equine cytochrome c protein sequence was used for cytochrome c and Saccharomyces Cerevisiae Database (yeastgenome.org, version 2006.10.06) was used for Dnm1G385D with a modification to include the Dnm1G385D sequence instead of Dnm1. Parent and fragment mass errors of 10 ppm and 0.4 Da were allowed. In the main search, potential oxidation of Met was allowed, and Cys was treated to be completely carbamidomethylated for Dnm1G385D and unmodified for cytochrome c. Lys modification of +138.06808 Da (intrapeptide link) and +156.07864 Da (monolink) are included in the first and the second round of the refinement search as potential modifications. Equine cytochrome c also includes protein N-terminus acetylation of +42.0106 Da and heme modification of +615.1700 Da in the refinement search. Reversed sequence was added in the search, and a log e value of −2 was used as a cut off. Half tryptic peptides were allowed in the refinement search but did not affect the final result and were disregarded.

In the SEQUEST, single protein sequences of equine cytochrome c and Dnm1G385D were used with a fragment tolerance of 1.0 Da. Modifications were the same as in X!Tandem: for both Dnm1G385D and cytochrome c, Met oxidation, Lys intrapeptide cross-link and monolink modifications; for

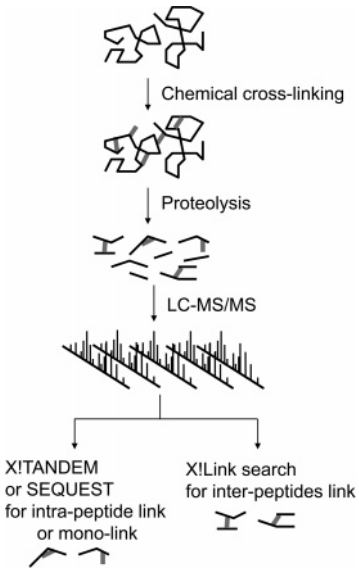
Dnm1G385D, Cys full carbamidomethylation; for cytochrome c, Cys heme modification and N-terminus acetylation. Fully tryptic peptides with Xcorr values higher than 2.0, 2.5, and 3.0 are accepted for +2, +3, and +4 or higher charge state peptide ions and a false peptide probability of 5% or higher is rejected. Peptide mass tolerance of 2.5 Da was used to include isotope errors in the search. Then, SEQUEST results were manually adjusted for the isotope errors (i.e., subtraction of 1.003355 Da for $\Delta m \approx +1$ Da; the mass difference between 12C and 13C) and mass errors higher than 10 ppm were rejected. All the intrapeptide cross-links were subjected to manual verification, whereas mono-links were confirmed only for those Xcorr lower than 2.5 and 3.5 for +2 and +3 or higher charged states, respectively, in SEQUEST and log e value higher than −4 in X!Tandem.

Development of X!Link. The homemade program, X!Link, was developed to analyze interpeptide cross-links. This program was written in Python 2.4 and currently does not support a graphic interface but works with text formatted data. It assumes primary amine reactive chemical cross-linking reagents such as BS3 and DSS (Disuccinimidyl suberate) are used, and thus allows cross-linking on primary amine sites, found either in Lys or in the unmodified N-terminus of a protein. Initially, it reads an information file that contains modified mass (i.e., 138.06808 Da for BS3), data file name, two output file names, two protein sequence file names, and other information such as filtering parameters. After it reads the protein sequences, it generates a list of tryptic peptide sequences and their corresponding masses. Only fully tryptic peptides were included that have at least one internal Lys or the N-terminus of the protein and a maximum of two missed cleavages was allowed. The current version is designed to read deisotoped MS/MS spectra in MGF format generated by MASCOT Distiller. X!Link automatically reads, analyzes, and reports each MS/MS spectrum one by one, so that there is no limitation in the number of MS/MS spectra. After it reads a MS/MS spectrum, it attempts to find a pair(s) of two peptides that has a determined cross-linked mass within the parent mass tolerance. If a pair is identified, X!link generates a fragment mass list of the cross-linked peptide and counts how many of them match with the experimental MS/MS spectrum. Once the two filtering criteria pass the preset threshold, the MS/MS score (the number of total fragments assigned divided by the number of total amino acids) and the number of major peaks assigned, the pair is considered to be correct and is reported to the output text files. There is also an additional optional filtering related to the parent mass error and the number of missed cleavages (see the following section).

In generating the fragments list, only b and y type fragments, their doubly charged forms, and triply charged forms (for +4 or higher charge state parent ions) are considered. Double fragmentation in a single peptide, often called an internal fragment, is rarely found in the CID of tryptic peptide ions using an ion trap mass spectrometer. However, double fragmentation at each side of a cross-linked peptide is often observed and its singly charged form is also included. Nomenclature by Schilling and co-workers was followed for these assignments:[14] i.e., $y_{3\alpha}b_{5\beta}$ for a cross-linked form between y_3 of peptide α and b_5 of peptide β.

The program generates two output files. The first one is a summary of each matching result: two peptide sequences, charge, parent mass error, number of miss-cleavages, number of total amino acids, number of total peaks matched, MS/MS score, and number of major peaks matched. The second file contains further information to help with manual validation: spectral number, a matched ion list with assignment, and a list of all the theoretical fragments. All matched results are manually evaluated to ensure correct assignment. Typical criteria for manual validation includes (1) at least four consecutive y or b type fragments, (2) all the major peaks assignable, (3) facile N-terminal and scarce C-terminal cleavage of Pro, (4) facile C-terminal cleavage of cross-linked Lys,[15] and (5) fragmentations from both peptides. The parameters used in this study are as follows: maximum peptide length for each peptide, 30; intensity threshold in MS/MS spectra, 10 count or 1% of base peak; parent mass tolerance, 10 ppm; fragment mass tolerance, 0.6 Da. MS/MS filtering parameters are optimized using statistical analysis as shown in the following section.

Scheme 1. Schematic for the Cross-Linking Experiment and Data Analysis



Results

Strategy for Cross-Linking Analysis. Scheme 1 shows the general outline of our experimental procedure and data analysis in which a chemical cross-linking step is added before a standard proteomics protocol. Purified protein or protein complexes are chemically cross-linked to form covalent bonds between two adjacent amino acids that contain functional groups reactive to the cross-linker: i.e., two Lys's within ∼20 Å. Cross-linked protein or protein complexes can be digested either in solution or in gel and analyzed by LC−MS/MS. Unlike others, we do not use labeling techniques or a comparison with non-cross-linked control samples to distinguish cross-linked peptides from other non-cross-linked ones. Hence, the success of this strategy strongly depends on accurate parent mass determination and MS/MS filtering to reject random matching.

For the identification of single peptides with intrapeptide cross-links or monolinks, we used established software such as X!Tandem or SEQUEST. Cross-linking is identified as a post-translational modification on a cross-linkable amino acid, Lys in this study, with a modified mass of (cross-linking part − 2H) for an intrapeptide cross-link or (cross-linking part − H + OH) for a monolink. In the case of an intrapeptide cross-link, even though there are two lysines that are cross-linked, X!Tandem or SEQUEST searches as if only one is modified. Therefore, we

Table 1. Intra- and Interpeptide Cross-Links Observed in Cytochrome c

cross-links[a]	distance[b], Å	number of MS/MS spectra	references[c]
Interpeptide cross-links			
5K-8K	<12	2	
5K-87K	16	9	10
7K-27K	21	48	14
8K-87K	17	8	
8K-100K	13	3	
39K-55K	<12	2	
39K-60K	<12	34	
39K-100K	18	1	
53K-79K	12	3	
55K-72K	13	1	
55K-73K	13	15	
60K-79K	20	7	
60K-86K	19	6	
72K-79K	<12	4	
72K-86K	16	13	32
72K-99K	23	3	
73K-86K	16	7	
73K-88K	16	1	
86K-88K	<12	6	
87K-100K	21	1	
88K-99K	17	6	
Total: 21		180	
Intrapeptide cross-links[d]			
7K-8K	<12	2	13,32
8K-13K	<12	1	
22K-25K	<12	17	
25K-27K	<12	28	10,31,32
53K-55K	<12	1	
55K-60K	13	2	
72K-73K	<12	12	
73K-79K	14	2	
86K-87K	<12	10	14
99K-100K	<12	5	14
Total: 10		77	

[a] See Supplementary Tables 1 and 2 for the detailed information about the identified cross-links. [b] Molecular distance is between the α carbons of the two cross-linked Lys's obtained from the RCSB Protein Data Bank using Jmol viewer (http://www.rcsb.org/, 1AKK.pdb, ref 33). [c] Cross-links observed by others. [d] For intrapeptide cross-links, X!Tandem and SEQUEST data were combined. The MS/MS spectra detected by both programs were counted only once.

needed to confirm the existence of two cross-linkable amino acids and that there was no fragmentation between the two.

For the identification of cross-links between two peptides, we developed an in-house program, X!Link. As Schilling and co-workers have classified,[14] there are many different kinds of interpeptide cross-links. Nonetheless, like others, we consider only a single cross-link between two peptides. Cross-linking experiments are usually performed under conditions that result in a minimum number of cross-links per protein to avoid structural distortion. Under these conditions, multiple cross-links between two or three peptides are supposed to be rare. The overall procedure is essentially the same with other shot gun analysis tools: parent mass filtering in combination with MS/MS scoring. We employ strict filtering, however, to minimize random matching: parent mass filtering in accurate mass and double MS/MS filtering. Filtering parameters were optimized and evaluated for false positive probability as described in a separate section. Final results were verified by manual inspection.

Cross-Links in Cytochrome c. A summary of the cross-links identified in cytochrome c using our approach is summarized in Table 1. A total of 21 interpeptide and 10 intrapeptide cross-links were identified in 257 MS/MS spectra. Detailed information on identified peptide ions can be found in Supplementary Tables 1 and 2 (see Supporting Information). On the basis of the NMR structure of cytochrome c (1AKK.pdb from http://www.rcsb.org/)[16], the α carbon distances between the cross-linked Lys residues that we identified are all within 24 Å, consistent with the length of a fully expanded BS3 and two Lys molecules as expected. Hence, under our experimental conditions, intramolecular cross-linking occurs without significant disturbance in molecular geometry.

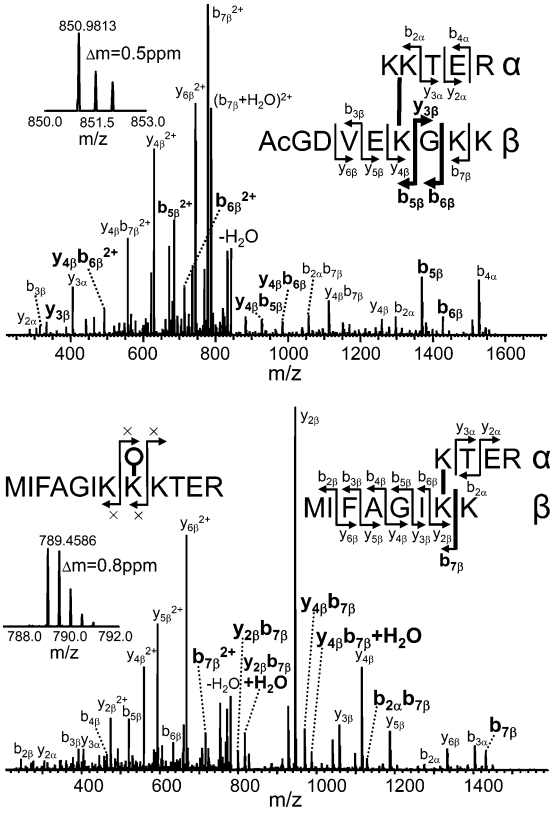


Figure 1. MS/MS spectra of interpeptide cross-links of ***KK***TER-AcGDVE**K**GKK (Top) and **K**TER-MIFAGI**K**K (Bottom) from cytochrome c. Inset mass spectra are a zoom scan of parent ions.

In some instances, the identification of the exact cross-linking sites was not straightforward and required careful manual inspection. Figure 1 shows two examples of such cases of interpeptide cross-links observed in cytochrome c. In the top panel, both the α and *β* peptides have two internal Lys's. Thus, further analysis was needed to determine which Lys pair was cross-linked. In such cases, cleavage between the two Lys's provides valuable information. In this case, we did not find clear evidence of which Lys in peptide α is cross-linked but we did observe $y_{3\beta}$, $b_{5\beta}$, $b_{6\beta}$, and their associated double fragments for peptide *β*. These fragments suggest that 5K of peptide *β* was the cross-linked site. Hence, both Lys's of peptide α were marked as bold and italic in Supplementary Table 1, ***KK***TER, whereas only 5K is marked as bold for the peptide *β*, AcGDVE**K**GKK, and this cross-link is assigned as "87K-5K or 88K-5K". A spectral counting in Table 1 was added to 87K-5K because 87K-5K is uniquely assigned elsewhere, **K**K-AcGDVE**K**GK, but 88K-5K was not detected in any other cross-links.

The bottom panel of Figure 1 shows another example of the complexities encountered in the assignment of cross-linking sites. The amino acid sequence KTER follows MIFAGIKK in the protein sequence and its cross-linked form has exactly the same

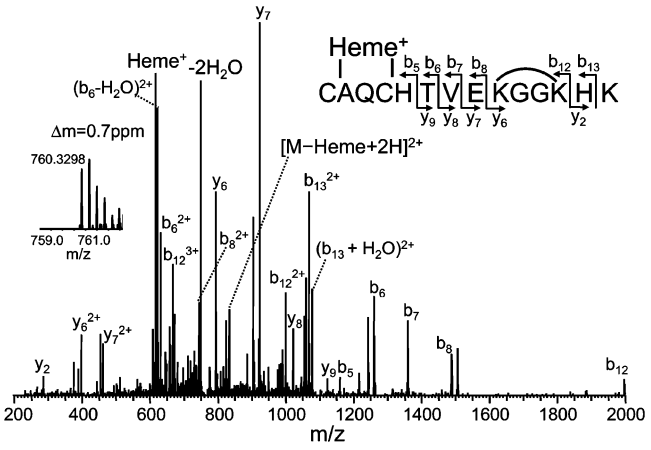


Figure 2. MS/MS spectrum of the peptide, CAQCHTVEKGGKHK, intrapeptide cross-linked by both heme and BS3. The heme group is cross-linked between the two Cys's and BS3 is cross-linked between the two internal Lys's. Inset mass spectrum is a zoom scan of the parent ion.

mass as the mono-linked MIFAGIKKKTER peptide. The MS/MS spectrum was detected as both an interpeptide cross-link by X!Link (Score = 2.83) and a monolink by X!Tandem (log e = −2.2). Among the three consecutive Lys's, X!Tandem suggested the second Lys as the monolink site. However, we could not find any evidence for the mono-linked peptide: absence of a unique cleavage product between the first and the third Lys, regardless of a monolink on any Lys. In contrast, $b_{7\beta}$ and the accompanying double fragments, $y_{2\beta}b_{7\beta}$ and $y_{4\beta}b_{7\beta}$, provided clear evidence to support the interpeptide cross-link. Hence, this spectrum was assigned as an interpeptide cross-link of **K**TER-MIFAGI**K**K. In another identified cross-link between the two consecutive peptide sequences, the monolink was subsequently invalidated after manual inspection: the cross-link GITWKEETLMEYLENP**K**K-YIPGT**K**MIFAGIK (Score = 1.16 and 1.42 for $Z = 4$ and 5, respectively) and the monolink GITWKEETLMEYLENPKKYIPGT*K*MIFAGIK (log e = −3.2 with $Z = 5$).

A c-type heme group is covalently bonded to two Cys's in the motif CXXCH (X = any amino acid). Therefore, it could be detected as a naturally occurring intrapeptide cross-link with the same approach used for the chemical cross-linking analysis. A similar strategy was adopted by Yang and co-workers to detect heme-attached peptides in *Shewanella oneidenis* cytochromes.[17] Their methodology used SEQUEST to search for a mass addition of +615.1694 Da, one hydrogen less than the mass of the heme, on Cys to take into account the fact that the heme group has formal charge of +1 in gas phase. To make this even more accurate, we used +615.1700 Da, one proton less than the mass of the heme, on Cys as a post-translational modification because SEQUEST or X!Tandem assumes a proton as a charge carrier, not hydrogen. Supplementary Table 2 shows all of the heme-attached peptides we identified. A total of 123 MS/MS spectra were observed to have heme as an intrapeptide cross-link and many of them also contain BS3 monolink or intrapeptide cross-link modifications.

Figure 2 shows a MS/MS spectrum of a peptide that contains both an attached heme and BS3 as intrapeptide cross-links. A heme fragment ion appeared as an intense peak at m/z 616.3 Da. The counter product of the heme fragment, $[M - \text{heme} + 2H]^{2+}$, was also prominent. As expected, any bond cleavage between the cross-linked amino acids, either between two Cys's or two Lys's, was not observed because the bond cleavage products are linked by the cross-linker and still have the same mass as the parent. It is also noteworthy that $b_{n-1}+H_2O$ fragments with its water taken from the C-terminus, such as $b_{13}+H_2O$ in Figure 2, are commonly observed in MS/MS spectra of cross-linked peptides due to its internal Lys.[18,19]

Table 2. Number of MS/MS Spectra Observed as a Cross-Linked Product for Each Lys of Cytochrome c[a]

	monolink	intrapeptide cross-link	interpeptide cross-link
5K	0	0	11
7K	0	2	48
8K	1	3	13
13K	22	1	0
22K	43	17	0
25K	3	45	0
27K	20	28	48
39K	16	0	37
53K	14	1	3
55K	3	3	18
60K	23	2	46
72K	12	12	21
73K	4	14	23
79K	4	2	14
86K	7	10	32
87K	3	10	18
88K	7	0	13
99K	4	2	10
100K	0	2	5

[a] For monolinks and intrapeptide cross-links, X!Tandem and SEQUEST data were combined. The MS/MS spectra detected by both programs were counted only once.

Table 2 summarizes the number of spectra observed as cross-linked products for each Lys of cytochrome c. We could detect all the Lys's as either intra- or interpeptide cross-links, which exemplifies the high sensitivity of our method. Some Lys sites appear to be favored for a monolink modification while other sites are favored for either intra- or interpeptide cross-links. For example, 22K is most abundantly detected as a monolink whereas 7K and 25K are most favored in inter- and intrapeptide cross-links, respectively.

Cross-Links in the Dnm1G385D Homodimer. We applied our methodology to study the interprotein interaction sites of the Dnm1G385D homodimer to demonstrate the widespread applicability of our method. Dnm1 is a dynamin-related GTPase that functions to mediate the division of mitochondria. Dynamin-related proteins (DRPs) are large self-assembling GTPases whose common function is to regulate membrane dynamics in a variety of cellular processes. DRPs contain three functionally important and distinct regions: a GTPase domain, a smaller middle region, and a C-terminal assembly region or GTPase effector domain.[20,21] These domains associate via intra- and intermolecular interactions to promote the self-assembly of dynamin into higher order filamentous and spiral-like structures and to stimulate GTP hydrolysis to a relatively high rate.[20−26] High-resolution structural data for full length dynamins are not available, likely because of the propensity of these proteins to self-assemble. To gain insight into the structural basis of self-assembly, we are probing of intra- and intermolecular interactions by identifying regions of close proximity through the determination of chemically formed peptide−peptide cross-links,

Dnm1G385D has been used for this analysis because it possesses a mutation in the middle region that blocks self-assembly and traps the protein as a stable homodimer, which

Table 3. Intra- and Interpeptide Cross-Links Observed in the Dimer Band of Dnm1G385D[a]

cross-links	number of MS/MS spectra
Interpeptide cross-links	
104K-111K	3
104K-128K	3
104K-300K	1
104K-748K	2
111K-128K	2
111K-137K	1
111K-285K	1
111K-318K	1
150K-156K	1
152K-193K	4
156K-193K	1
237K-331K	3
237K-336K	1
285K-300K	6
293K-300K	5
310K-318K	5
331K-338K	3
336K-747K	3
393K-480K	1
393K-549K or 627K	1
491K-715K	2
491K-718K	2
549K-556K	3
556K-605K	1
556K-613K	1
603K-616K	1
605K-615K	1
605K-616K	2
613K-616K	3
613K-626K	2
615K-626K	3
741K-748K	3
549K or 627K-600K	1
549K or 627K-605K	2
549K or 627K-615K	1
549K or 627K-616K	1
549K or 627K-626K	1
549K or 627K-674K	1
Total: 38	79
Intrapeptide cross-links	
150K-152K	2
152K-156K	1
285K-293K	2
603K-605K	2
613K-615K	3
615K-616K	2
626K-627K	1
715K-718K	6
Total: 8	19

[a] See Supplementary Tables 3 and 4 for the detailed information on identified peptides.

is the fundamental building block for assembly.[13] Table 3 shows a summary of the cross-links obtained in the SDS-PAGE resolved dimer of Dnm1G385D cross-linked by BS3. We could detect a total of 46 cross-links, 38 interpeptide and 8 intrapeptide cross-links, in 98 MS/MS spectra. This is, to our knowledge, the largest number of cross-links ever identified in a single experiment or single protein complex to date, demonstrating the utility and sensitivity of our method. Detailed information on the identified peptide ions can be found in Supplementary Tables 3 and 4 (see Supporting Information). The tryptic peptide KR exists twice in the protein sequence, 549K and 627K, and could not be differentiated. Hence, as seen in Table 3, the cross-linking site was assigned as "549K or 627K".

Manual interpretation of each cross-link was not always straightforward, however. For example, Dnm1G385D has two

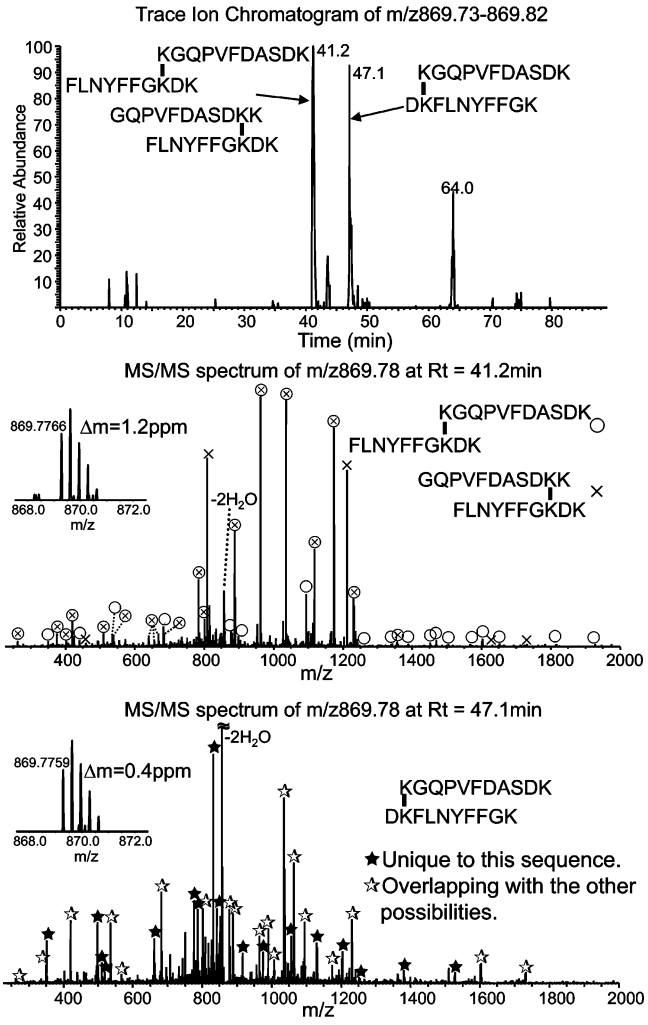


Figure 3. (Top) Trace ion chromatogram of the Dnm1G385D homodimer in the mass range of *m*/*z* 869.73−869.82. (Middle) MS/MS spectrum of *m*/*z* 869.78 at the retention time 41.7 min corresponding to the mixture of triply charged cross-links of **K**GQPVFDASDK-FLNYFFG**K**DK and GQPVFDASD**K**K-FLNYFFG**K**DK. (Bottom) MS/MS spectrum of *m*/*z* 869.78 at the retention time 47.1 min corresponding to the triply charged cross-link of **K**GQPVFDASDK-D**K**FLNYFFGK.

pairs of Lys miscleaved tryptic peptides with very similar sequence and exactly the same mass: DKFLNYFFGK and FLNYFFGKDK; KGQPVFDASDK and GQPVFDASDKK. Their combination generates four different cross-links that are very similar in sequence and exactly the same in parent mass: D**K**FLNYFFGK-**K**GQPVFDASDK, D**K**FLNYFFGK-GQPVFDASD-**K**K, FLNYFFG**K**DK-**K**GQPVFDASDK, FLNYFFG**K**DK-GQPVFDASD**K**K. In addition, these two peptides are consecutive in sequence and the possible monolink of the FLNYFFG*K*D-*KK*GQPVFDASDK peptide (one of three internal Lys is monolinked) also has exactly the same mass. We found that three of the cross-links actually exist as shown in Figure 3. The top panel shows a trace ion chromatogram in the *m*/*z* range of 869.73−869.82 Da of the LC−MS/MS data, *m*/*z* range around the triply charged cross-links. Two chromatographic peaks at the retention times of 41.2 and 47.1 min were observed as the aforementioned cross-links in the X!Link search and are shown in the middle and bottom panel MS/MS spectra. X!Link reported all four cross-links for the first peak and three of them for the second. After careful manual inspection, the first peak at the

retention time of 41.2 min was assigned as a mixture of **K**GQPVFDASDK-FLNYFFG**K**DK and GQPVFDASD**K**K-FLNYFFG**K**DK as shown in the middle panel. Many peaks were overlapping between the two, and are marked as ⊗, but there are some peaks that could be uniquely assigned to either of the two, and are marked as ○ or × (see Supplementary Table 5 for the fragments assignment, Supporting Information). We did not find any fragment uniquely assignable to the other possibilities. The second chromatographic peak at the retention time of 47.1 min was assigned as **K**GQPVFDASDK-D**K**FLNYFFGK in the MS/MS spectrum shown at the bottom panel. In spite of the many fragments assignable not only to this cross-link, but also to the other possibilities, marked as ☆, some fragments were unique to this cross-link, and are marked as ★, allowing us to make this assignment (Supplementary Table 5). X!Tandem misassigned the same spectrum as the monolink of FLNYFFG*K*DKKGQPVFDASDK (log e = −5.0), which was invalidated by manual inspection. This example demonstrates the importance of careful manual inspection for the final interpretation of data, especially when several cross-links exist with a similar sequence.

We looked for other interpeptide cross-links in Dnm1G385D that are consecutive in sequence and could also be interpreted as a monolink. There were two more cases: D**K**LPDIK-T**K**LNTLISQTEQELAR and D*K*LPDI*KT*K*L*NTLISQTEQELAR; **K**GQPVFDASDK-**K**R and *K*GQPVFDASD*KK*R. X!Link picked out both cases as interpeptide cross-links while X!Tandem reported only the former as a monolink. After manual inspection, both cases turned out to be interpeptide cross-links and thus the monolink assignment was invalidated.

A total of 42 Lys sites of the Dnm1G385D homodimer were modified by BS3, as either inter- or intrapeptide cross-links or monolinks. That corresponds to a Lys coverage of 72% among a total of 58 Lys's in the Dnm1G385D molecule. Most of the Lys's not covered in this study are expected to be either unexposed to the surface or to be found within tryptic peptides too long to analyze. For example, a monolink modification on 95K will result in a 46 residue tryptic peptide, 64R-109R, or even a larger peptide if involved in an interpeptide cross-link and, therefore, most likely it will not be extracted from the gel and/or detected by LC−MS/MS.

Statistical Analysis of False Positives to Optimize Filtering Parameters. The success of our strategy strongly depends on our ability to minimize incorrect assignments. For X!Tandem and SEQUEST, which we used in the monolink and intrapeptide cross-link search, many publications are available about the statistical aspects of random matching.[27−30] Here, we simulated false positive distributions of X!Link as a function of various input parameters to find optimal input parameters that minimize false positive assignments. They are compared with the average true positive distribution obtained from the above two data sets.

To get statistics on random matching, we needed to search X!Link against a large number of peptides. However, unlike other shotgun tools, X!Link was not designed to search against multiple protein sequences but designed to search only one or two protein sequence(s) that produce a limited number of tryptic peptides. Hence, a huge pseudo protein sequence was generated by combining all equine protein sequences (equine protein database supplied by Thermo for BioWorks 3.2, 346 total protein entries, a total of 75 453 amino acid residues) back-to-back to make a single sequence. LC−MS/MS data obtained from a control gel band (no cross-linking) and the dimer band of Dnm1G385D were searched against this pseudo whole-equine sequence. This pseudo sequence generates 60.8 times more cross-linkable tryptic peptides than Dnm1G385D (8333 and 137 unique tryptic peptides for the pseudo sequence and Dnm1G385D, respectively, with at least one internal Lys and up to two missed cleavages), but the random matching is estimated to be 3700 times greater because X!Link looks for a combination of two peptides and the number of possible cross-links increases as the square of the number of peptides. Accordingly, the false positive data obtained above were scaled by 3700. The true positive data were obtained by combining the two data sets from cytochrome c and Dnm1G385D and scaled by two.

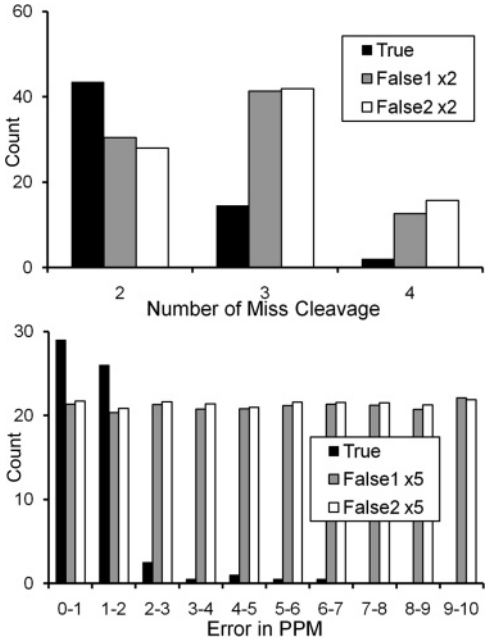


Figure 4. True and false positive distributions of interpeptide cross-linked ions as a function of the number of missed cleavages (Top) and parent mass error (Bottom). True positives are the average of the two data sets from cytochrome c and the Dnm1G385D homodimer band. False positives are simulated by searching the two LC−MS/MS data set, control (no cross-linking; False1) and dimer band (False2) of Dnm1G385D, against a pseudo whole-equine sequence.

Figure 4 shows the distribution of true and false positives as the number of miss cleavages and parent mass error. Because of the internal Lys residues in each peptide for cross-linking, interpeptide cross-links have at least two missed cleavages. The number of true positives decreased significantly for missed cleavage while the number of false positives decreased slowly with a random population of approximately 1:2:1 for missed cleavages of two, three, and four, because up to two missed cleavages were allowed for each peptide. As the bottom panel shows, parent mass errors for the true positives were mostly within 3 ppm (root-mean-square average of 1.5 ppm), whereas false positives were randomly spread. Mass error of 4 ppm or higher was attributed to insufficient ion counts or significant contamination in the FT Zoom scan. Here, we made a simple optional filtering that removes four missed cleavages with more than 3 ppm mass error and three missed cleavages with more than 5 ppm mass error. When we applied this filtering, we removed 35∼37% of false positives, while having only a 0.6% of chance of losing the true positives. Significantly, we did not lose any true positives in the current data set using this filtering.

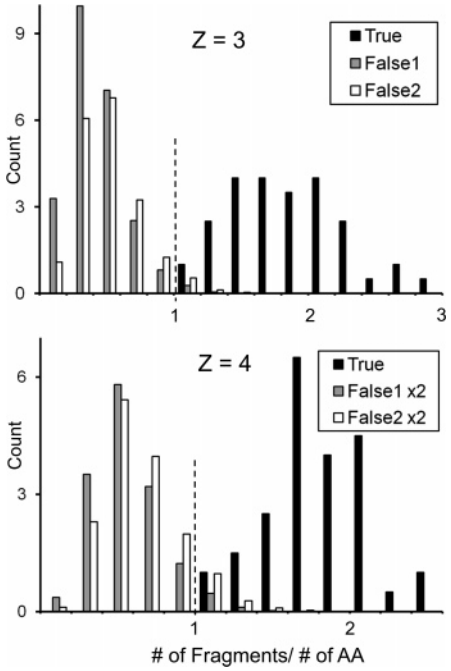


Figure 5. True and false positive distributions of interpeptide cross-linked ions as a function of MS/MS score, defined as the number of fragments divided by the number of total amino acids, at $Z = 3$ (Top) and $Z = 4$ (Bottom). MS/MS score of 1 is suggested as cut off value for both $Z = 3$ and 4.

Filtering in the MS/MS level was conducted by using the number of fragment ions assigned divided by the total number of amino acids as MS/MS score. We attempted a few different scoring systems including the percentage of assigned fragments,[9] but we did not improve the results further. Figure 5 shows the true and false positive distributions for $Z = 3$ and $Z = 4$ as a function of the MS/MS score. The true positives were well distinguished from the false positives. If we used a minimum score of 1.0 for both $Z = 3$ and $Z = 4$, most of the false positives were removed while maintaining all the true positives. Similar distributions were acquired for $Z = 2$ and Z > 4 and a minimum score of 1.6 and 0.8 were suggested, respectively.

Additional MS/MS filtering was applied by counting assignable major peaks among the top ten intensities. This filtering concept came from the fact that most of the major peaks are assigned in manual validation. Figure 6 shows the population of true and false positives as a function of the number of major peaks assigned. We observed a clear bimodal distribution of true positives centered at 6 or 7 and false positives centered at 0. When the number of major peaks assigned was three or above, the number of false positives rapidly diminished. We observed a single true positive with only one major peak assigned (5+ ion of **K**YIPGTK-TGQAPGFTYTDAN***K***N***K***GITWK) and two true positives with only two major peaks assigned (5+ ion of **K**HPVYR-SQQDIQLN**K**TVEESLDKEEDYFR, 4+ ion of K**K**TER-**K**IFVQK). From careful manual inspection, we observed the low number of major peaks assigned in these cross-linked peptide ions originated from 4+ charged fragment ions that were not considered in the algorithm. The same cross-linked sequences were observed in other charge states with more than three major peaks assigned. Therefore, their exclusion did not affect the sensitivity of the method and thus a minimum of three major peaks assigned is recommended. Even though the minimum cut off was three in the computer algorithm, all the major peaks were assignable in the manual validation.

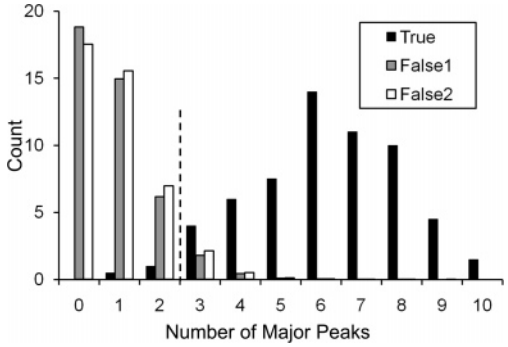


Figure 6. True and false positive distributions of interpeptide cross-linked ions as a function of the number of major peaks assigned by X!Link. Major peaks are defined as the top ten intensity peaks. A minimum number of three is suggested as a cut off value.

Applying all the optimized filtering parameters discussed in this section to the Dnm1G385D homodimer band data, a false discovery rate of 1.0% was calculated from the scaled number of cross-links obtained for the pseudo protein sequence divided by the number for the correct protein sequence. In addition, when we tried the X!Link search on control samples (no cross-linking) against the correct protein sequences or cross-linked samples against a few incorrect protein sequences, no result was reported by X!Link demonstrating the effectiveness of the filtering used.

Discussion

In spite of its promise for protein structural analysis, mass spectrometry-based identification of protein cross-links has been limited by its lack of high-sensitivity data analysis tools. Under the bold assumption that identification of cross-linked peptides is possible by direct mass spectrometry analysis alone, we developed a high throughput and high sensitivity computer algorithm for the analysis of protein cross-linking by mass spectrometry, which we term *shotgun cross-linking analysis*. The success of this method largely depends on accurate parent mass determination and high quality MS/MS acquisition. LTQ-FT, a hybrid mass spectrometer we employed for the current study, provides accurate mass information with back-end FT mass spectrometer and high quality MS/MS data with front-end linear ion trap, LTQ. Indeed, when low-resolution LC−MS/MS data obtained with LTQ was analyzed instead, we failed to identify cross-links due to too many false positives (data not shown).

As a part of our method, we adopted the programs X!Tandem and SEQUEST for identifying monolinked peptides and intrapeptide cross-links because they are well-established for identifying modifications on single peptides. Monolinks could be treated as usual post-translational modifications. In contrast, a significant difference exists between intrapeptide cross-links searched using the above programs and actual cross-linked peptides. These programs assume modification on only one residue of the cross-linked sites: i.e., MIFAGIK*KK or MIFAGIKK*K (K*: +138.06808Da on K) for MIFAGI**KK**K (cross-link between two bold Lys's). Hence, they allow fragments induced by bond breakage between the two Lys's, which are not observed in the actual intrapeptide cross-link as shown in Figure 2.

Compared to intrapeptide cross-links, the identification of interpeptide cross-links is far more challenging. First, random matching increases exponentially with the number of theoretical tryptic peptides. Thus, the search should be limited to the minimum protein sequence(s). We are currently searching against only one protein sequence for a single protein or homodimer and two protein sequences for heterodimer. It will be an even greater challenge, however, to apply this technique to multi-protein complexes. Second, because of the many basic sites found within the cross-linked peptides, multiply charged ions are predominant making the assignment of cross-links more difficult. Significantly, 87% of interpeptide cross-links we found in this study are from peptide ions of triply or higher charged states (see Supplementary Tables 1 and 3). We found cross-links with charge states up to +6. Thus, it is essential to have accurate mass and charge information to filter out most of the random matching. The interpeptide cross-link search algorithm, X!Link, was designed with these difficulties in mind. X!Link was optimized to report a minimum number of false positives without compromising the identification of true positives, especially in multiply charged cross-links. Even though this program is designed for the direct analysis of cross-linked data set, it could also be used to analyze MS/MS spectra predetermined to be cross-linked peptides by comparing with control sample or isotopic labeling.

Cytochrome c is often used to test new methods of cross-linking analysis as it is small protein with its entire lysines surface exposed and available for chemical cross-linking. With our approach, we identified all the Lys's as a cross-linked product and detected naturally occurring heme cross-linked peptides in 123 MS/MS spectra. Additionally, among all of the Lys pairs that are within 20 Å, 33% were identified as either intra- or interpeptide cross-links in a single LC−MS/MS data set (27 pairs out of a total of 83 pairs). This is a very high rate of identification considering all of the Lys pairs will not essentially be cross-linked due to the preference in the reaction kinetics.

In the tandem mass spectrometric approach to identify cross-linking sites in cytochrome c, Schilling and co-workers detected an interpeptide cross-link (7K-27K) and two intrapeptide cross-links (39K-53K, 86K-87K).[14] In an accurate mass mapping approach without MS/MS, Dihazi and co-workers identified an interpeptide cross-link (5K-88K) and five intrapeptide cross-links (25K-27K, 39K-53K, 86K-87K, 99K-100K, 5K-7K, or 5K-8K or 7K-8K) using three cross-linkers (Sulfo-DST, BS3, Sulfo-EGS).[31] Pearson and co-workers adopted an isotopic labeling and LC−MS/MS approach and identified an interpeptide cross-link (72K-86K) and four intrapeptide cross-links (7K-8K, 8K-13K, 25K-27K, 39K-53K).[32] We found all these cross-links except the intrapeptide cross-link of 39K-53K in 39K-55K sequence (**K**TGQAPGFTYTDAN**K**NK). This cross-link generates mostly internal fragments rather than y or b type fragments and is difficult to identify by traditional tools such as SEQUEST and X!Tandem. This suggests a limitation of the current approach for intrapeptide cross-links with two Lys's separated by many internal amino acids. We also could not find unique evidence for 5K-88K as discussed previously (top panel of Figure 1).

Recently, Seebacher and co-workers developed a LC−MALDI tandem mass spectrometric approach combined with isotope-coded cross-linkers and successfully applied this technique to the Colicin E7 DNAse/Im7 heterodimer and other proteins. In their application to cytochrome c, they identified 12 cross-links using DSG and DSS. Their identification of 4 cross-links with one side on the protein N-terminus is questionable considering equine cytochrome c has an acetylated N-terminus. It is possible though that those 4 cross-links may have come from unacetylated protein existing in a low amount (private communication). Among the 8 other cross-links, we could find only one interpeptide cross-link (5K-87K) and one intrapeptide cross-link (25K-27K). But, other than 39K-53K and 87K-88K, 3 others have α carbon distance of more than 24 Å and we could find 29 other cross-links within 24 Å.

The analysis of the chemical cross-links in the Dnm1G385D homodimer demonstrates the high sensitivity and widespread applicability of our method. Analysis of the protein homodimers is very challenging because we cannot differentiate inter*protein* cross-links from intra*protein* cross-links. Mixing heavy and light isotope-labeled subunits would be an ideal approach to distinguish inter- and intraprotein cross-links in homomultimer complex.[5] In a more simplified approach, we are currently trying to differentiate intraprotein cross-links from interprotein cross-links by analyzing cross-links in the Dnm1G385D monomer band (data not shown).

In addition to high throughput and high efficiency programming, careful manual validation, which is crucial for any cross-linking analysis using mass spectrometry, is an essential part of our method. It is the current bottleneck to faster data processing. We plan to develop a graphic interface to facilitate this step. In many cases, manual validation is straight forward and can be easily accomplished once we have the graphic interface developed. However, as seen in Figures 1 and 3, some cross-links need to be carefully verified when multiple assignments are possible.

Conclusion

We developed a new approach for cross-linking analysis and successfully applied the methodology to detect inter- and intrapeptide cross-links and monolinks from cytochrome c and the Dnm1G385D homodimer. This methodology produces few false positives while identifying most of the true positives. We believe this method is among the most sensitive and easy-to-use methods developed so far for cross-linking analysis, owing mostly to the simplistic procedure and computer algorithm. Most importantly, it simply needs a set of LC−MS/MS data acquired with a high-resolution mass spectrometer; it does not need a comparison with a non-cross-linked control sample or the use of costly stable isotopes. We are in the process of developing this program further to include a graphic interface, options to use various cross-linkers and enzymes and to identify/include various amino acids modifications, and a probabilistic score.

Acknowledgment. J.N. is supported by NIH grants 1R01EY015924 and 5R01GM062942. L.L. is supported by an NIH postdoctoral fellowship 1F32GM078749-01.

Supporting Information Available: Supplementary Tables 1 and 3 are lists of all the interpeptide cross-linked ions in cytochrome c and the Dnm1G385D homodimer cross-linked by BS3. Supplementary Tables 2 and 4 are lists of all the intrapeptide cross-links and monolinks observed in cytochrome c and the Dnm1G385D homodimer. This material is available free at http://pubs.acs.org.

References

(1) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5802−5806.

(2) Huang, B. X.; Kim, H. Interdomain Conformational Changes in Akt Activation Revealed by Chemical Cross-linking and Tandem Mass Spectrometry. *Mol. Cell. Proteomics* **2006**, *5*, 1045−1053.

(3) Ihling, C.; Schmidt, A.; Kalkhof, S.; Schulz, D. M.; Stingl, C; Mechtler, K.; Haack, M.; Beck-Sickinger, A. G.; Cooper, D. M. F.; Sinz, A. Isotope-labeled cross-linkers and Fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1100−1113.

(4) Back, J. W.; Notenboom, V.; Koning, L. J.; Muijsers, A. O.; Sixma, T. K.; Koster, C. G.; Jong, L. Z. Identification of cross-linked peptides for protein interaction studies using mass spectrometry and O-18 labeling. *Anal. Chem.* **2002**, *74*, 4417−4422.

(5) Taverner, T.; Hall, N. E.; O'Hair, R. A. J.; Simpson, R. J. Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry. *J. Biol. Chem.* **2002**, *277* 46487−46492.

(6) Sinz, A.; Waug, K. Mapping Protein Interfaces with a Fluorogenic Cross-Linker and Mass Spectrometry: Application to Nebulin-Calmodulin Complexes. *Biochemistry* **2001**, *40*, 7903−7913.

(7) Tang, Y.; Chen, Y.; Lichti, C. F.; Hall, R. A.; Raney, K. D.; Jennings, S. F. CLPM: A cross-linked peptide mapping algorithm for mass spectrometric analysis. *BMC Bioinformatics* **2005**, *6*, S9.

(8) Chu, F. X.; Shan, S. O.; Moustakas, D. T.; Alber, F.; Egea, P. F.; Stroud, R. M.; Walter, P.; Burlingame, A. L. Unraveling the interface of signal recognition particle and its receptor by using chemical cross-linking and tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16454−16459.

(9) Gao, Q. X.; Xue, S.; Doneanu, C. E.; Shaffer, S. A.; Goodlett, D. R.; Nelson, S. D. Pro-CrossLink. Software tool for protein cross-linking and mass spectrometry. *Anal. Chem.* **2006**, *78*, 2145−2149.

(10) Seebacher, J.; Mallick, P.; Zhang, N.; Eddes, J. S.; Aebersold, R.; Gelb, M. H. Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing. *J. Proteome Res.* **2006**, *5*, 2270−2282.

(11) Craig, R.; Beavis, R. C. TANDEM: matching proteins with mass spectra. *Bioinformatics* **2004**, *20* 1466−1467.

(12) Eng, J. K.; McCormack, A. L.; Yates, J. R., III. An approach to correlate MS/MS data to amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.

(13) Ingerman, E.; Perkins, E. M.; Marino, M.; Mears, J. A.; McCaffery, J. M.; Hinshaw, J. E.; Nunnari, J. Dnm1 forms spirals that are structurally tailored to fit mitochondria. *J. Cell Biol.* **2005**, *170*, 1021.

(14) Schilling. B.; Row, R. H.; Gibson, B. W.; Guo, X.; Young, M. M. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 834−850.

(15) Gaucher, S. P.; Hadi, M. Z.; Young, M. M. Influence of Crosslinker Identity and Position on Gas-Phase Dissociation of Lys−Lys Crosslinked Peptides. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 395−405.

(16) Banci, L.; Bertini, I.; Gray, H. B.; Luchinat, C.; Reddig, T.; Rosato, A.; Turano, P. Solution structure of oxidized horse heart cytochrome c. *Biochemistry* **1997**, *36*, 9867−9877.

(17) Yang, F.; Bogdanov, B.; Strittmatter, E. F.; Vilkov, A. N.; Gritsenko, M.; Shi, L.; Elias, D. A.; Ni, S. S.; Romine, M.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D. Characterization of purified c-type heme-containing peptides and identification of c-type heme-attachment sites in Shewanella oneidenis cytochromes using mass spectrometry. *J. Proteome Res.* **2005**, *4*, 846−854.

(18) She, Y.-M.; Krokhin, O.; Spicer, V.; Loboda, A.; Garland, G.; Ens, W.; Standing, K. G.; Westmore, J. B. Formation of $(b_{n-1} + H_2O)$ ions by collisional activation of MALDI-formed peptide $[M + H]^+$ ions in a QqTOF mass spectrometer. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1024−1037.

(19) Thorne, G. C.; Ballard, K. D.; Gaskell, S. J. Metastable Decomposition of Peptide $[M + H]^+$ Ions via Rearrangement Involving Loss of the C-Terminal Amino Acid Residue. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 249−257.

(20) van der Bliek, A. M. Functional diversity in the dynamin family. *Trends Cell Biol.* **1999**, *9*, 96−102.

(21) Muhlberg, A. B.; Warnock, D. E.; Schmid, S. L. Domain structure and intramolecular regulation of dynamin GTPase. *EMBO J.* **1997**, *16*, 6676−6683.

(22) Marks, B.; Stowell, M. H. B.; Vallis, Y.; Mills, I. G.; Gibson, A.; Hopkins, C. R.; McMahon, H. T. GTPase activity of dynamin and resulting conformational change are esstential for endocytosis. *Nature* **2001**, *410*, 231−234.

(23) Sever, S.; Muhlberg, A. B.; Schmid, S. L. Impairment of dynamin's GAP domain stimulates receptor-mediated endocytosis. *Nature (London)* **1999**, *398*, 481−486.

(24) Smirnova, E.; Shurland, D.-L.; Newman-Smith, E. D.; Pishvaee, B.; van der Bliek, A. M. A model for dynamin self-assembly based on binding between three different protein domains. *J. Biol. Chem.* **1999**, *274*, 14942−14947.

(25) Warnock, D. E.; Hinshaw, J. E.; Schmid, S. L. Dynamin self-assembly stimulates its GTPase activity. *J. Biol. Chem.* **1996**, *271*, 22310−22314.

(26) Zhu, P. P.; Patterson, A.; Stadler, J.; Seeburg, D. P.; Sheng, M.; Blackstone, C. Intra- and intermolecular domain interactions of the C-terminal GTPase effector domain of the multimeric dynamin-like GTPase Drp1. *J. Biol. Chem.* **2004**, *279*, 35967−35974.

(27) Tabb, D. L.; Smith, L. L.; Breci, L. A.; Wysocki, V. H.; Lin, D.; Yates, J. R., III. Statistical Characterization of Ion Trap Tandem Mass Spectra from Doubly Charged Tryptic Peptides. *Anal. Chem.* **2003**, *75*, 1155−1163.

(28) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75*, 4646−4658.

(29) Eriksson, J.; Fenyö, D. Probity: A Protein Identification Algorithm with Accurate Assignment of the Statistical Significance of the Results. *J. Proteome Res.* **2004**, *3*, 32−36.

(30) Fenyö, D.; Beavis, R. C. A Method for Assessing the Statistical Significance of Mass Spectrometry-Based Protein Identifications Using General Scoring Schemes. *Anal. Chem.* **2003**, *75*, 768−774.

(31) Dihazi, G. H.; Sinz, A. Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2005−2014.

(32) 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.

(33) Banci, L.; Bertini, I.; Gray, H. B.; Luchinat, C.; Reddig, T.; Rosato, A.; Turano, P. Solution structure of oxidized horse heart cytochrome c. *Biochemistry* **1997**, *36*, 9867−9877.

PR070234I