Finding Chimeras: a Bioinformatics Strategy for Identification of Cross-linked Peptides*

Feixia Chu‡, Peter R. Baker, Alma L. Burlingame, and Robert J. Chalkley§

Chemical cross-linking, followed by identification of the cross-linked residues, is a powerful approach to probe the topologies and interacting surfaces of protein assemblies. In this work, we demonstrate a new bioinformatics approach using multiple program modules within the software package “Protein Prospector” that greatly facilitates the discovery of cross-linked peptides in chemical cross-linking studies. Examples are given for how this approach has been used for defining interfaces in heterodimeric and homodimeric protein complexes, both of which provide results in close agreement with crystal structures, verifying the reliability of the approach. Molecular & Cellular Proteomics 9:25–31, 2010.

Proteins act in the context of dynamic protein complexes and interaction networks (1, 2). To gain mechanistic insights into various cellular processes, it is important to devise enabling strategies for the mapping of protein interaction surfaces. Chemical cross-linking has been an established method to study protein interaction partners for decades (3) and has been recently revived by conjunctional mass spectrometric analysis of the cross-linking reaction products (4). Mass spectrometric identification of the cross-linked residues provides valuable spatial restraints in defining protein interacting surfaces. Nevertheless, comprehensive analysis of proteolytic digests of cross-linked protein assemblies is challenging. The difficulty is largely due to the fact that most generic cross-linking reagents react with multiple residues and meanwhile undergo competing side reactions (with solvent, for example), inactivating one of the reactive groups. This means that the dominant reaction products are where only one end of the reagent reacts with a peptide, and these are referred to as dead-end modified peptides (5). To facilitate the detection and separation of cross-linked species, various moieties have been introduced into the cross-linking reagents including fluorophores (6), isotope tags (7, 8), cleavable sites (9), affinity handles (10–14), and a benzyl marker for tandem mass spectrometric analysis (15). However, mere incorporation of these moieties cannot distinguish the cross-linked peptides from dead-end modified products (16).

The high complexity of these samples requires tandem mass spectrometric fragmentation data for confident identification of the cross-linked peptides. Although highly desirable, available bioinformatics tools for automatic analysis of fragmentation spectra of the cross-linked samples are far from robust, mainly limited by the difficulty in identifying fragments that contain components from cross-linked peptide species. Thus far, most of the reported software either rely on pre-screening for cross-linker-containing peptides by their distinct isotope pattern (17–19) or require manual entry of peak lists for each predetermined cross-linked peptide candidate (5). Another strategy has just been reported where mass modification searching allowed identification of cross-linked candidates (20). It is based on selecting quadruply charged precursors as potentially being cross-linked peptides and then searching the obtained spectra one at a time to try to determine whether they are cross-linked peptides and identify the sequences involved. Here we present an alternative approach using available programs in the bioinformatics software package “Protein Prospector” (21, 22), which allows analysis of all MS/MS spectra acquired in a data set in one analysis. Thus, one can identify unmodified and cross-linked peptides together in a single search result. In all cross-linked protein complexes tested, our new strategy has led to the discovery of not only all cross-linked peptides previously identified through manual analysis but new cross-linked peptides as well. These comprehensive results were achieved with much higher throughput, demonstrating the robust and comprehensive applicability of the approach to elucidating structural information about protein assemblies.

EXPERIMENTAL PROCEDURES

All methods connected to the study of the Ffh-FtsY complex are described in a previous publication (23).

Materials—Disuccinimidyl suberate (DSS) was purchased from Pierce. Sequencing grade modified porcine trypsin was from Pierce. The protease-deficient Escherichia coli strain 27C7 was used for the expression of Gly-Gly-His (GGH) N-terminally tagged ecotin D137Y as described previously (24).

Cross-linking Reactions and Mass Spectrometric Analysis—Cross-linking reactions, SDS-PAGE separation, in-gel digestion, and LC-MS analyses were carried out as described previously (25). Essentially, after SDS-PAGE separation of the cross-linked complex, both dimer and monomer bands were sliced and digested. Peptides were extracted, and the extraction solution was dried down to 10 µl. For differential LC-MS analysis, a 1-µl aliquot of the digestion mixture

---

1 The abbreviation used is: DSS, disuccinimidyl suberate.
Identifying Cross-linked Peptides

RESULTS

Recently, we developed an unspecified mass modification search algorithm in Batch-Tag, a program in the Protein Prospector package, to identify small unexpected peptide modifications (22). We perceived that it could also be a powerful tool for identifying cross-linked peptides by treating the cross-linked peptide moiety as a single peptide with a large modification of unspecified mass (27) (the schematic work flow is shown in Fig. 1). After proteolytic digestion, peptides of a cross-linked protein complex are analyzed by LC-MS/MS. Components of the protein complex are first identified on the basis of unmodified peptides by conventional protein data-base searching methods. This list of proteins is then used as a restricted database for an unspecified mass modification search, looking for peptides with large modifications, which are potentially cross-linked peptide products. Searches can be performed against a target-decoy database to obtain a measure of reliability for results, for example by searching against a concatenated database of a target database and a sequence-randomized version of the same database (a sequence-reversed database can also be created using Protein Prospector scoring was performed as published previously (21) and peptide N-terminal glutamine conversion to pyroglutamate. Considered were methionine oxidation, protein N-terminal acetylation, and 300 ppm, respectively, were used, and the only modifications allowed. Precursor and fragment mass tolerances of 50 and 300 ppm, respectively, were used, and the only modifications considered were methionine oxidation, protein N-terminal acetylation, and peptide N-terminal glutamine conversion to pyroglutamate. Protein Prospector scoring was performed as published previously (21) at a flow rate of −300 nL/min. The HPLC eluent was connected directly to the microion electrospray source of a QSTAR Pulsar QqTOF mass spectrometer (Applied Biosystem/MDS Sciex, Foster City, CA).

Mass Spectrometric Data Analysis Using Batch-Tag and MS-Bridge—LC-MS data were acquired using the Analyst QS software (Applied Biosystems, Foster City, CA). Peak lists were created using the Mascot.dll (version 1.6b20) and then searched using Batch-Tag within Protein Prospector. Data were initially searched against the full Swiss-Prot database (downloaded April 24, 2008 with 320,363 entries) to which the sequences of tagged ecotin, Ffh, and FtsY were added. Fully tryptic peptides were required with up to three missed cleavages allowed. Precursor and fragment mass tolerances of 50 and 300 ppm, respectively, were used, and the only modifications considered were methionine oxidation, protein N-terminal acetylation, and peptide N-terminal glutamine conversion to pyroglutamate. Protein Prospector scoring was performed as published previously (21) where a given score is credited for each fragment match with the score weighting dependent on the ion type matched; e.g. a y ion scores more than an internal ion. These scores were then converted to expectation values as described previously (22) by determining the distribution of scores for random answers and calculating a probability and then expectation value of a given score being in this distribution. The list of accession numbers for all identified proteins in the sample was then used for a second search. In this search, DSS (or bis(sulfosuccinimidyl) adipate) modification to lysines and the protein N terminus was also considered. Because of errors in the automated peak list generation, a mass shift of +1 Da that is not considered on any fragment ions was permitted to allow matching of spectra where the second isotope was incorrectly assigned as the monoisotopic peak. As well as these defined variable modifications, this search also allowed for a single mass modification of between −100 and +4000 Da to lysine residues or the peptide N terminus. Fragment mass tolerance was increased to 0.25 Da because average masses must be used to predict the mass modifications because of a lack of knowledge of the elemental composition of an unknown modification. Other search parameters were the same as for the initial search. Initial acceptance criteria required a reported expectation value of less than 0.1. All results were then manually verified.

Spectra that were confidently matched to peptides with a large mass modification were then queried using MS-Bridge, another program in Protein Prospector. The input to MS-Bridge consisted of the masses of cross-linked peptides observed, the sequences of the cross-linked protein(s), enzymes used in digestion, amino acid(s) participating in cross-linking reactions, elemental composition of the linker bridge, and any variable modifications to be considered. The program returns a list of potential unmodified, modified, or cross-linked peptide combinations that correspond to the experimentally observed masses input. Having already identified one of the cross-linked peptides from the Batch-Tag search, this step will suggest identities for the second cross-linked peptide, and the second sequence can be input into the MS-Product display of the Batch-Tag results to show peak matches to the two peptides in the same spectrum (see supplemental Fig. 1 for an example). There is no scoring associated with the matching of this second peptide, so the user must make an objective judgment about whether they believe the second peptide assignment based on the number of possibilities for the identity of the second peptide returned by MS-Bridge and how many unexplained peaks from the assignment of the first peptide are now explained as fragments from the second peptide.
Prospector (28)). After identification of the first peptide moiety (Fig. 1, peptide A) in a cross-linked complex, the Protein Prospector program MS-Bridge (23, 25) is able to report all possible cross-linked peptide combinations from the proteins in the restricted database that produce the correct molecular mass product. The other cross-linked peptide component (Fig. 1, peptide B) is determined based on knowing the sequence of the first peptide moiety (Fig. 1, peptide A) identified from the Batch-Tag search. The fragmentation spectrum of the cross-linked species is then matched against in silico fragmentation of both peptide components using MS-Product, which allows displaying matches to multiple sequences (see supplemental Fig. 1 for an example). This allows confirmation of identification of the cross-linked product.

To assess sensitivity (ability to identify as many components present as possible) and specificity (as high a percentage as possible of the answers reported as significant are correct) of this new strategy, we first applied it to a previously characterized system, the heterodimeric complex of Ffh-FtsY. The specific interaction between Ffh and FtsY mediates cotranslational protein targeting of bacterial secretory and membrane proteins to the plasma membrane (29). We previously carried out a chemical cross-linking study of the Ffh-FtsY complex and manually identified cross-linked peptides that were only present in the digests of the cross-linked complex (23). We hereby reanalyzed the *Thermus aquaticus* Ffh-FtsY data set using the new bioinformatics approach. In an unspecified mass modification search, we identified 22 peptides with large mass modifications corresponding to a second peptide attached by a cross-link (supplemental Table 1). There were no matches to the decoy part of the concatenated peptide attached, so the first glycine is 3 residues before the start of the protein sequence in the database, Gly(3)Lys9 (the protein has an N-terminal tag of Gly-Gly-His extension and Asp137 to Tyr mutation was used for this study. Cross-linked homodimer was separated from unlinked monomer by SDS-PAGE, and both samples were processed through the same analytical procedures, including in-gel digestion and LC-MS/MS analysis, followed by mass modification searching and MS-Bridge analysis. Multiple cross-linked species were retrieved after mass modification searching in Batch-Tag; however, it is difficult to distinguish intersubunit from intrasubunit cross-links at this stage. Hence, we compared the identified peptides between analyses of the dimer and monomer samples in Search Compare, which led to the identification of five intersubunit cross-links (supplemental Table 1).

As an example, the Batch-Tag mass modification search reported a species (m/z 670.8541) as the peptide Ala\(^{132}\)Arg\(^{142}\) with a mass addition of 1387 amu on Lys\(^{135}\). MS-Bridge then generated a list of all possible cross-linked peptide combinations from the protein that match to the mass of this cross-linked species. Knowledge that one component was the peptide Ala\(^{132}\)–Arg\(^{142}\), identified in the Batch-Tag mass modification search, allowed the N-terminal peptide Gly\(^{1–3}\)-Lys\(^9\) (the protein has an N-terminal tag of Gly-Gly-His attached, so the first glycine is 3 residues before the start of the protein sequence in the database, Gly\(^{1–3}\)) to be unequivocally identified as the second peptide component. This identification was further confirmed by fragment ions from the Gly\(^{1–3}\)-Lys\(^9\) peptide moiety (Fig. 2a). Besides the normal cross-links, we observed a species 56 amu lower in mass than the cross-linked Gly\(^{1–3}\)-Lys\(^9\) and Ala\(^{132}\)–Arg\(^{142}\) dipeptide. Batch-Tag identified extensive fragmentation ions from both peptide moieties. Fragment ions containing the cross-linker moiety (667.3227+ and 708.8927+; Fig. 2a) are 56 amu lower than corresponding ions (695.3422+ and 736.8832+; Fig. 2b) in the normal cross-linked peptide. Thus, the mass difference is likely due to the cross-linker itself (Fig. 2b). Conceivably, low level contamination of butyrate (instead of suberate) linker has led to the formation of this cross-linked species. Although this result did not provide additional information on the ecotin dimer, it exemplifies the power of this unbiased mass modification searching strategy. One would predict that there may be low levels of other cross-linked products formed from this butyrate linker. However, no others were identified in the database searching, although there is no guarantee that
these components were selected for MS/MS analysis. The intersubunit cross-links are mapped onto the x-ray structure of wild type ecotin in Fig. 2c. The cross-linking results are in good agreement with both the crystal structure and a previous study on this mutant (24).

Chemical cross-linking studies to reveal protein interaction surfaces are most commonly carried out on binary complexes such as those in this study. However, the approach demonstrated here should be applicable to the analysis of larger complexes and more complicated mixtures. To get an idea of how well this bioinformatics approach performs in a more complex sample background, the peak lists from the Ffh-FtsY complex data set were combined with a peak list from a publicly available standard protein mixture data set (26), also acquired on a QSTAR mass spectrometer (QS20060131_S_18mix_02). This sample contained just under 40 proteins, so the analysis of the combined data set provides assessment of the ability of this bioinformatics approach to accurately identify the cross-linked species in the presence of many non-cross-linked proteins and peptides. An initial search of this data set without considering mass modifications returned a list of 39 identified proteins, including Ffh and FtsY. Large mass modification searching in Batch-Tag was then carried out on the combined data set against a concatenated database of these 39 proteins and randomized versions of each.

Analysis of the results showed that the spectra identified as cross-linked species in previous analysis against only Ffh and FtsY were practically all identified again to the same peptide and mass modification. However, the expectation values assigned were in the range of 10–30-fold more conservative. Thus, some reported expectation values to correct results were greater than 1. However, the false positive rate is low even at very high expectation values. For example, with a maximum expectation value threshold set at 20, a total of 57 spectra with mass modifications greater than 500 Da were reported (see supplemental Table 2). Of these, 49 were matched to either Ffh or FtsY, and all of these matches were to spectra in the cross-linking data sets rather than spectra from the standard protein mixture despite the fact that there were twice as many spectra in the standard data set as the cross-linking data (1721 versus 916). Upon closer analysis, for three of these eight matches to proteins other than Ffh or FtsY, part of the correct peptide was reported, but the wrong monoisotopic peak and/or charge was supplied to the search engine: in the cases of the alkaline phosphatase and myoglobin peptide matches, when given the correct information these spectra were both matched to unmodified versions of peptides containing the same sequence, and the spectrum at m/z 927.733+ was actually a 6+ precursor that matched an extended version of the same tryptic peptide as reported in the mass modification search but also contained monomethylated and dimethylated lysine residues (chemical modifications introduced into the trypsin to try to reduce autolysis). These results suggest that the software is still fairly effective at separating real hits from random results but that the absolute values for the expectation value calculations have become meaningless. This problem is caused by the conversion of probabilities into expectation values that is performed by multiplying the probability of a precursor of the correct mass obtaining a given score by the number of precursors in the database that have the correct precursor mass. If any modification over a large mass range is considered, then the number of theoretical peptides considered is excessively large in comparison with the number of real possibilities, so the expectation value estimation becomes unnecessarily conservative. For example, when identifying the interprotein cross-linked product at m/z 554.30+1, when searching for mass modifications against only the two cross-linked protein sequences Batch-Tag considered 1368 potential peptide precursors, whereas in the 39-protein database search, 56,948 precursors were considered, which is actually a 4 times greater value than the number of unmodified tryptic peptides of the correct mass in the whole Swiss-Prot database. However, this same explosion in numbers of precursors considered is the same for all spectra whether confidently matched or not, so it represents a fairly constant shift in expectation values for all results whether correct or not. Hence, the use of a target-decoy database searching strategy and then selection of a suitable expectation value threshold based on these results should still be an effective strategy.

Having identified cross-linked peptide spectra, the next step in the workflow is the use of MS-Bridge to try to predict the second peptide present in the cross-linked complex. The presence of extra proteins only had a marginal effect on the ability to identify the second peptide: for the six interprotein cross-linked peptides identified, for three of them MS-Bridge reported that there was still only one possible second peptide, whereas in the 39-protein database search, 56,948 precursors were considered, which is actually a 4 times greater value than the number of unmodified tryptic peptides of the correct mass in the whole Swiss-Prot database. However, this same explosion in numbers of precursors considered is the same for all spectra whether confidently matched or not, so it represents a fairly constant shift in expectation values for all results whether correct or not. Hence, the use of a target-decoy database searching strategy and then selection of a suitable expectation value threshold based on these results should still be an effective strategy.

In this case, the second peptide was confident and could be confirmed by mass spectrometry. However, this is not the case for the second peptide present in the cross-linked complex. The Batch-Tag unspecified mass modification searching step is necessary to identify the second peptide present in the cross-linked complex. Herein, we have developed a bioinformatics approach for the in-depth analysis of cross-linked protein complexes. The Batch-Tag unspecified mass modification searching step is

**Fig. 2.** a, structure and low energy CID spectrum of the cross-linked species with an m/z value of 670.854+ from cross-linked ecotin dimer digest. The fragments containing the cross-linker bridge are colored in green. b, structure and low energy CID spectrum of the cross-linked species with an m/z value of 656.844+ from cross-linked ecotin dimer digest. The fragments containing the cross-linker bridge are colored in green. c, crystal structure of ecotin. The intersubunit cross-linked residues are represented as magenta (in chain 1) or red (in chain 2) sticks, and the distances between cross-linked residues are displayed as dotted lines.
Identifying Cross-linked Peptides

comprehensive and unbiased and considers all types of peptides present, including unmodified peptides, dead-end modified peptides, and intrapeptide and interpeptide cross-linked products. In addition, the fragment ions that contain cross-linked side chains are scored while identifying the cross-linked peptide components. Therefore, our approach retrieves "chimeric" cross-linked peptides with higher sensitivity and specificity than current methods that only consider unmodified fragment ions. However, it currently does not consider fragment ions that contain the cross-linker but include fragments of both peptide chains (5). Assignment of these ion types will be a future development of the software.

This strategy has allowed the identification of a large number of cross-linked peptides in essentially all protein assemblies we have tested thus far. This level of comprehensiveness in data analysis is important because the identification of multiple cross-links within a complex offers valuable distance constraints for downstream computational modeling of protein interaction surfaces. Furthermore, Batch-Tag mass modification searching is suitable for any cross-linker, including the study of natively disulfide-linked peptides, obviating dependence on isotope-coded cross-linkers and additional sample complexity introduced by these cross-linkers. The two studies presented here were of purified binary protein complexes, but we also simulated a more complex mixture and showed that the bioinformatics strategy still identifies cross-linked products in a more complex background. The major reason why this approach scales well with complexity is because, despite the fact that the number of potential cross-linked combinations grows exponentially with increasing numbers of proteins, our strategy identifies one-half of the complex at a time, so the increase in possibilities at each step is only a linear increase. The occasions where the increase in protein candidates has the most deleterious effect are when one of the peptides in the cross-linked complex is very short (2–4 amino acids). Sequences of this length can occur in many proteins, so even if it is possible to determine the sequence of the cross-linked peptide, it may not be possible to assign it to a particular protein. Because improving the mass accuracy of measurement of the cross-linked peptides correlates with a decreasing number of possible cross-linking combinations reported in MS-Bridge, high resolution and mass accuracy data will be beneficial for large protein complexes.

The presented strategy does not currently score the match to the second peptide, which would be highly beneficial for producing a measure of reliability. However, with our mass modification strategy it would be possible to perform a subsequent search where having identified one of the peptides we specify a modification of exactly the mass corresponding to the first peptide, which would then allow the second peptide to be identified in a database searching strategy. This is a feature we are seeking to develop in a future version of the software.

A challenge of all cross-linking studies is that the products of interest are highly substoichiometric. The results presented here, like most bioinformatics approaches, were dependent on the precursors of interest being selected for fragmentation. In our simulation of a more complex data set, we increased the number of non-cross-linked peptide spectra, but we did not simulate the problem that if there are more unmodified peptides in the mixture then the chances of the cross-linked peptides of interest being selected for fragmentation analysis are reduced. It is possible to bias the selection of precursors for fragmentation toward cross-linked products by only fragmenting precursors of higher charge state (most cross-linked tryptic peptide products will have two free amino groups and a basic residue in each peptide, so they will most commonly be quadruply charged, whereas very few unmodified tryptic peptides will have this high charge state). However, an enrichment strategy for modified peptides by incorporating a tag into the cross-linker would be the most effective approach (10–14).

In summary, our new bioinformatics approach and software tools represent a significant step forward in the ability to extract maximal information from chemical cross-linking studies and will hopefully encourage wider use of this type of strategy for characterizing protein complex structures. The Protein Prospector software presented here is publicly available on the web at http://prospector.ucsf.edu.

Acknowledgments—We thank Sami Mahrus and Charles S. Craik at the University of California San Francisco for providing ecotin mutant, GGH-ecotin D137Y, for this study.

* This work was supported, in whole or in part, by National Institutes of Health Grants NCRR 01614, NCRR 12961, and NCRR 15804 (all to A. L. B.) from the National Center for Research Resources.

** The on-line version of this article (available at http://www.mcponline.org) contains supplemental Fig. 1 and Tables 1 and 2.

‡ Present address: Dept. of Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, NH 03824.

§ To whom correspondence should be addressed: Mass Spectrometry Facility, Dept. of Pharmaceutical Chemistry, University of California, Box 2240, 600 16th St., Rm. GH-N474A, San Francisco, CA 94143-2240. Tel.: 415-476-5189; Fax: 415-502-1655; E-mail: chalkley@cgl.ucsf.edu.

REFERENCES

1. Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dümpelfeld, B., Edelmann, A., Heurter, M. A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A. M., Scheider, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J. M., Kuster, B., Bork, P., Russell, R. B., and Superti-Furga, G. (2006) Proteome survey reveals modularity of the yeast cell machinery. Nature 440, 631–636

2. Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrín-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Lalev, A., MenA, F., Wong, P., Starostine, A., Canete, M. M., Vlasblom, J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rillstone, J. J., Gandi, K., Thompson, N. J., Musso, G., St Onge, P., Ghanny, S., Lam, M. H., Butland, G., Altf-AUl, A. M., Karaya, S., Shilatifard, A., O’Shea, E.,

30 Molecular & Cellular Proteomics 9.1
Identifying Cross-linked Peptides

Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440**, 637–643.

Wong, S. S. (1991) *Chemistry of Protein Conjugation and Cross-linking*. CRC Press, Boca Raton, FL.

Sin, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* **25**, 663–682.

Chu, F., Mahrus, S., Craik, C. S., and Burlingame, A. L. (2006) Isotope-coded and affinity-tagged cross-linking (ICATXL): an efficient strategy to probe protein interaction surfaces. *J. Am. Chem. Soc.* **128**, 10362–10363.

Wlodawer, A., and Chang, C. Y. (2002) Three-dimensional protein structures and their applications. *Nature Struct. Biol.* **9**, 488–494.

Singh, P., Shaffer, S. A., Scherl, A., Holman, C., Pluetznzer, R. A., Larson Freeman, T. J., Miller, S. I., Hernandez, P., Appel, R. D., and Goodlett, D. R. (2008) Characterization of protein cross-links via mass spectrometry and an open-modification search strategy. *Anal. Chem.* **80**, 8799–8806.

Chalkley, R. J., Baker, P. R., Huang, L., Hansen, K. C., Allen, N. P., Rexach, M., and Burlingame, A. L. (2005) Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer: II. new developments in Protein Prospector allow for reliable and comprehensive automatic analysis of large datasets. *Mol. Cell. Proteomics* **4**, 1194–1204.

Chalkley, R. J., Baker, P. R., Medzhiradskszy, K. F., Lynn, A. J., and Burlingame, A. L. (2008) In-depth analysis of tandem mass spectrometry data from disparate instrument types. *Mol. Cell. Proteomics* **7**, 2386–2398.

Chu, F., Shan, S. O., Moustakas, D. T., Alber, F., Egea, P. F., Stroud, R. M., Walter, P., and Burlingame, A. L. (2004) 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.* **101**, 16454–16459.

Chu, F., Shan, S. O., Moustakas, D. T., Alber, F., Egea, P. F., Stroud, R. M., Walter, P., and Burlingame, A. L. (2004) Identification of novel quaternary domain interactions in the Hsp90 chaperone, GRP94. *Protein Sci.* **13**, 1260–1269.

Chakraborty, T., and Swaminathan, S. (2006) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214.