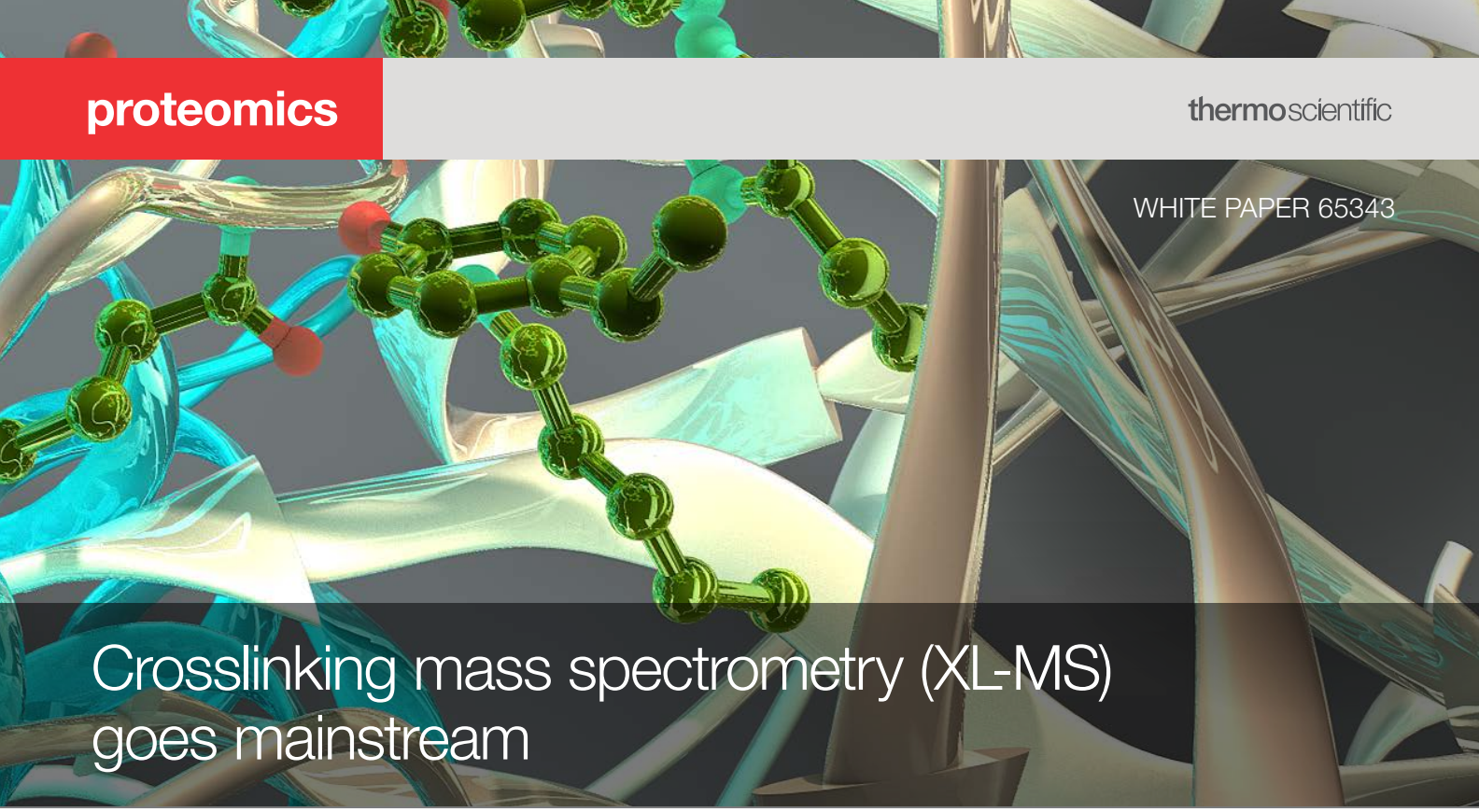


WHITE PAPER 65343

Crosslinking mass spectrometry (XL-MS) goes mainstream

Mass spectrometrists and structural biologists guide to using XL-MS to understand the structure and function of molecular machines

Author

Julian Saba, Thermo Fisher Scientific, San Jose, CA, USA

Keywords

Crosslinking mass spectrometry, XL-MS, protein-protein interactions, structural biology, protein complexes, crosslinking reagents (DSSO, DSBU), MS-cleavable crosslinker, ETD/EThcD, XlinkX, Proteome Discoverer software

Executive summary

Examining protein-protein interactions is crucial to understanding the roles proteins play in a biological system. The current techniques used to study protein interactions include X-ray crystallography, cryo-electron microscopy and nuclear magnetic resonance, all of which require significant protein amounts for analysis. There has been talk about mass spectrometry for many years and of crosslinking as a bona fide application to study protein-protein interactions. But, will this technique and application ever cross the chasm and become one that is adopted as mainstream in terms of a workflow, and ease of use?

Crosslinking mass spectrometry has emerged as a powerful tool for studying protein-protein interactions, especially now as there are fit for purpose reagents, advanced mass spectrometry technologies for intelligent analysis and specialized software packages designed for crosslinking studies. This white paper will focus on the challenges of performing crosslinking analyses, and how development of the latest tools from sample preparation through to data analysis have made this workflow more informative, streamlined and standardized.

ThermoFisher SCIENTIFIC

Introduction

To understand the roles that proteins play in biological systems, researchers must examine the functions proteins perform. While many proteins carry out their functions independently, the majority of proteins interact with one another for biological activity. Therefore, proteins should be studied in the context of protein-protein interactions, to fully understand their functions. Current techniques used to study protein interactions include X-ray crystallography, cryo-electron microscopy (cryo-EM) and nuclear magnetic resonance (NMR). However, these approaches require significant amounts of highly purified proteins and may not allow for the analysis of proteins in their native conditions. Furthermore, many proteins are simply not amenable to these types of analysis, thereby limiting the accessibility of these techniques.

Mass spectrometry (MS) techniques have been utilized in the research community for many years to study protein structure and protein-protein interactions, albeit usually with highly specialized research groups. However, due to their complexity and requirement for specialized sample preparation, advanced mass spectrometry feature requirements as well as fit-for-purpose data analysis, they have lagged behind general proteomics analyses.

The white paper presented here is intended to introduce crosslinking mass spectrometry (XL-MS), and how it has traversed from a manual, non-standardized research tool, emerging as a powerful solution for studying protein-protein interactions. This document will address the advantage of this technique as it relates to traditional approaches, as well as the complimentary role it plays in concert along-side these techniques. The latest tools and workflows that have been developed for XL-MS will be discussed. A special emphasis will be placed on new crosslinking reagents, novel mass spectrometers and software for data processing.

Protein interaction studies

Proteins are central to cellular function, playing crucial roles in nearly every biological process that occurs within a cell, from gene expression to cell growth and proliferation, intercellular communication and apoptosis. Examining the roles proteins play in biological processes can be challenging due to their dynamic nature and their characteristics within the cell. Cells do not exist in a vacuum isolated from outside forces. They are continuously stimulated by external factors that change their dynamics and properties, which in turn can affect the proteins inside the cell. Additionally, not all cells are identical, and since proteins are expressed in a cell type-dependent manner, proteins will vary depending upon the cell being examined. Furthermore, proteins that are used to complete specific tasks may not always be expressed or activated within a cell. These protein characteristics suggest a complexity that can be difficult to investigate, particularly when trying to understand protein function in a biological context. The complexity is further exacerbated by the fact that vast majority of proteins interact with one another for biological activity.

Protein interactions can be classified as either permanent or transient. Permanent interactions result in the formation of strong protein complexes. However, one of the biggest challenges to examining protein-protein interactions is that most interactions are transient. These interactions are weak, occurring only for a brief period, as part of a single cascade or other metabolic function within cells. Transient interactions control the majority of the cellular processes, including protein modification, transport, folding, signaling, and cell cycling. The examination of transient interactions is challenging due to lack of techniques available that measure interactions in real time.

Chemical crosslinking reagents provide a means for capturing protein-protein complexes by covalently binding them together as they interact. More recently, crosslinking combined with mass spectrometry has emerged as a powerful technique for investigating protein-protein interactions. In XL-MS, chemical crosslinkers are used to chemically join components of interacting complexes. This is followed by MS analysis of the joined complexes, enabling *in vivo* and *in vitro* approaches to study protein-protein interactions while maintaining the original interacting complex. The ability to visualize the interacting regions lets researchers create distance maps within the protein complexes or within the protein itself allowing for the generation of low resolution three-dimensional maps of these interactions.

There are several advantages to performing XL-MS over more traditional techniques:

- **Sample size:** Small sample size is required, typically in the range of nanogram (ng) for MS analysis. In comparison, NMR analysis requires sample amounts in the milligram (mg) range. Additionally, MS analysis can be performed directly at the proteome-wide level, something that is impossible to achieve with X-ray crystallography or NMR.
- **Protein purity:** XL-MS does not require highly pure protein for analysis compared to cryo-EM, X-ray crystallography or NMR.
  - In X-ray crystallography, it is often very difficult to produce well defracting crystals needed for analysis. This is further exacerbated in complexes where more than one protein needs to be crystalized. The purity and the order of the protein plays an important role in crystallization.
  - In cryo-EM, the images that are generated contain a low signal-to-noise ratio. The resulting images are very noisy with little contrast. Data analysis requires picking of the particles, associating particles into structurally identical groups, and then averaging grouped particles into a high resolution three dimensional structure. When the proteins are pure, less mistakes are made during the grouping process.
  - In NMR-based analysis, purity of the protein contributes to the quality of the data. For example, non-specific protein interactions can change the chemical shifts of the protein or protein complex of interest. This may culminate in changes all over the complex and to a mix of chemical shifts. The resulting mix of chemical shifts associated with a particular expected resonance, might be “diluted” by all possible signals that arise from the different binding interactions.
- **Analysis time:** XL-MS analysis can also be done in a shorter time relative to other techniques. In X-ray crystallography, the ability to generate diffracting crystals can be downright challenging. Crystal formation may take as short as an hour or as long as months to form. For NMR, the limiting step can be data acquisition. It may take several experiments to get the minimum amount of data to solve the structure. For XL-MS, once the sample is prepped, the data acquisition and data processing is very straightforward. The entire experiment can be completed in few hours.
- **Examination of interactions close to the physiological state of an organism:** XL-MS has the ability to do generate interaction information that is biologically relevant. Traditional approaches are limited by which proteins can be readily expressed or crystalized. The protein-protein interactions must be strong enough to survive general sample preparation steps of extraction and purification. As such, all of these issues make it very challenging to examine interactions in their native state.
- **Complement nature to other MS techniques for structure analysis:** XL-MS is complementary to other MS techniques such as hydrogen deuterium exchange mass spectrometry (HDX-MS), such that HDX-MS can be used make informed decisions about the regions to be studied by XL-MS. For example, HDX-MS is routinely used to examine the conformational flexibility of protein complexes while XL-MS provides information on distances between interacting regions. If the interacting regions are in perpetual motion, acquiring XL-MS information will not be very useful. Therefore HDX-MS can be used to make informed decisions about the regions studied by XL-MS. Furthermore, XL-MS can fill in information that might be missed by traditional approaches.
- **Easy addition to labs already with MS:** The XL-MS workflow can be performed by any proteomics laboratory as no special equipment or setup are required with the exception chemical reagents.

XL-MS is not just limited to protein-protein interaction studies; it can also be used to elucidate protein or protein complex structures. Protein structural information can be obtained independently or simultaneously, as part of protein interactions studies. For protein structural analysis, XL-MS is used in parallel with high-resolution techniques such as cryo-EM or X-ray crystallography to obtain structural information such as protein complexes, multi-subunit complexes and protein stoichiometry. Here, XL-MS provides distance constraints between regions within the protein, creating low-resolution, three-dimensional structural information or a general topology of the protein’s structure. Finally, XL-MS can also be used to study protein interactions with small molecules, nucleic acids, lipids and recently to study proteome wide interactions.

Crosslinking-mass spectrometry workflow

The primary workflow for XL-MS is very similar to a bottom-up proteomics workflow (Figure 1).

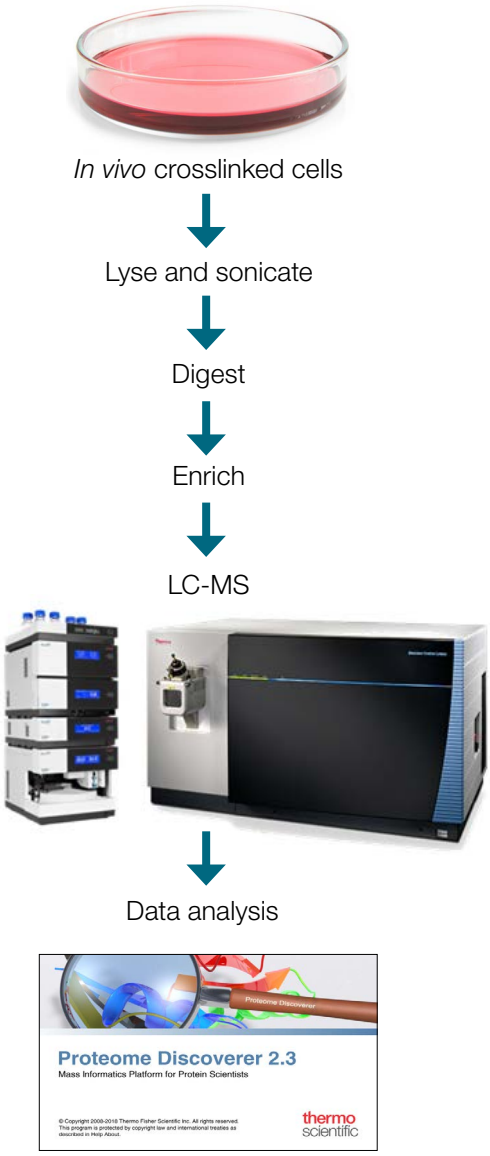


Figure 1. Schematic representation of a XL-MS workflow.

First, crosslinking reagents are used to covalently link interacting proteins or peptides that are in close proximity because of their interaction. If the crosslinking is done at the protein level, then the samples are digested to peptides with an appropriate enzyme. An enrichment step is incorporated upon digestion to isolate crosslinked peptides (Figure 2).

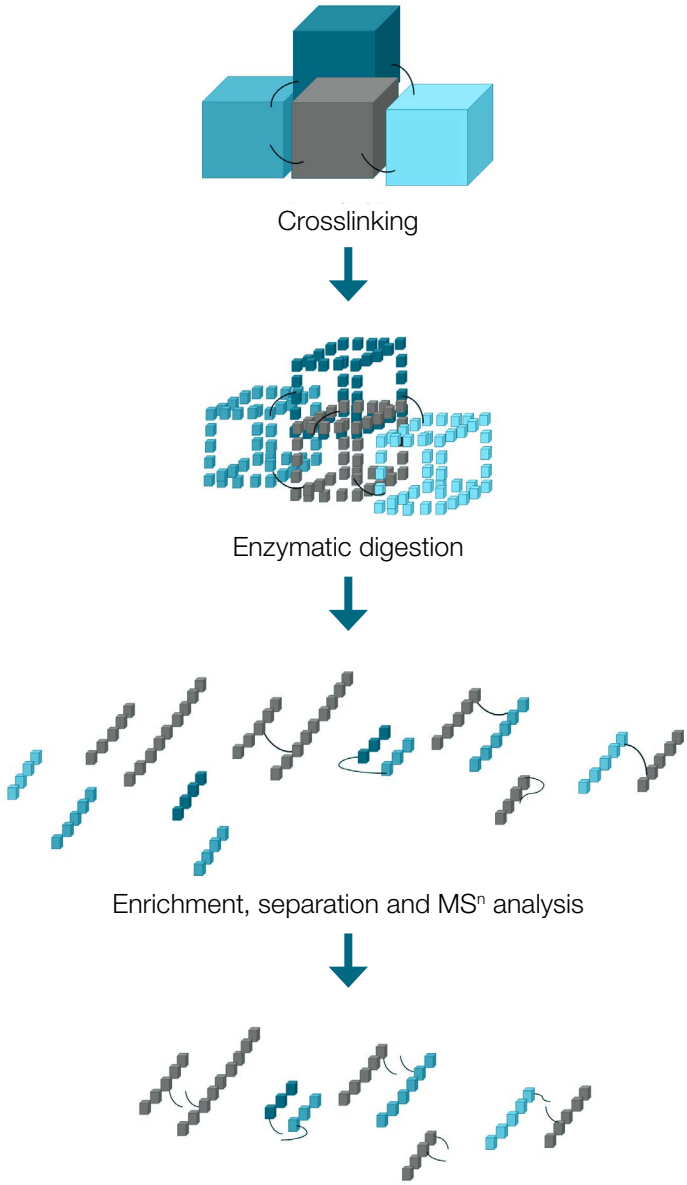


Figure 2. Simplified representation of sample preparation for XL-MS.

The samples are then separated and introduced into the mass spectrometer for MS^2 analysis. Specifically designed software is used for data interpretation. Currently, XL-MS has evolved to a very mature approach that can be implemented in any proteomics laboratory. The following sections will discuss these in more detail.

Chemical crosslinkers

Overview of crosslinker structure and chemistry

Whether it is examining protein-protein interactions or protein structures, chemically connecting different regions in a single protein or across different proteins by covalent bonds is required prior to MS analysis. Crosslinking reagents are used to achieve these covalent bonds. These reagents contain reactive ends to specific functional groups (primary amines, sulfhydryls, etc.) on proteins/peptides backbone or side chains. When these reagents are used to crosslink two regions on a single protein, it results in an intramolecular crosslink that stabilizes the protein's tertiary or quaternary structure. The crosslinking of regions between two different proteins results in intermolecular crosslinks that stabilize protein-protein interactions. The idea behind this approach is that by fixing the distance between interacting regions, researchers can use this distance information to reconstruct three-dimensional maps or structures of individual proteins or the protein-protein interactions using the distance constraints provided by the known length of the crosslinker.

In its simplest form, crosslinkers are composed of two reactive functional groups separated by a spacer arm. The spacer arm is the molecular span of the crosslinker, the region that separates the two functional groups. It is needed to compensate for the steric effects that dictate the distance between potential reaction sites for crosslinking. Typically, short spacer arms are used in intramolecular crosslinking and longer spacer arms for intermolecular crosslinking studies.

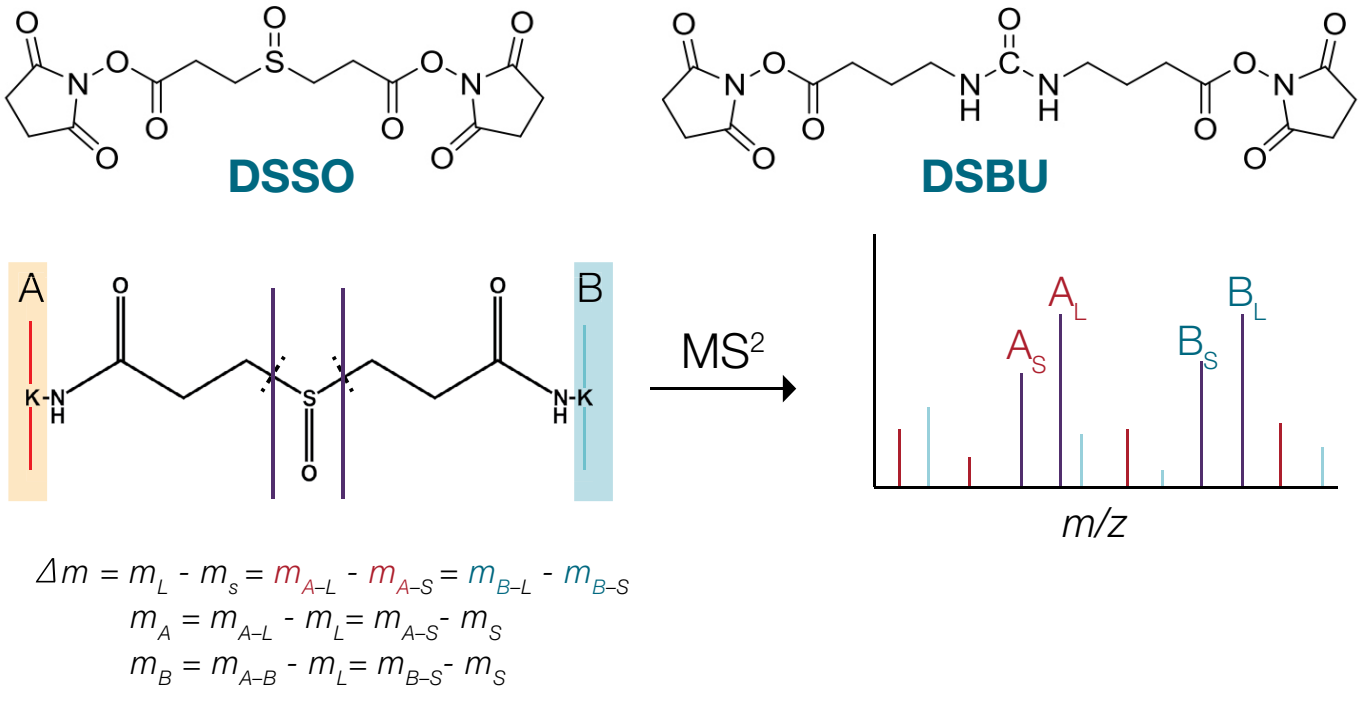
Crosslinkers can be further classified as homo-bifunctional or hetero-bifunctional. Homo-bifunctional crosslinkers have the same reactive groups at both ends of the spacer arm. Crosslinking with these types of linkers occurs in a one step process. This is ideal for capturing a "snapshot" of all protein interactions. Hetero-bifunctional crosslinkers have different reactive groups on either end of the spacer arm. These reagents not only allow for single-step conjugation of molecules that have the respective target functional groups, but they also allow for sequential (two-step) conjugations that minimize undesirable polymerization or self-conjugation.

Crosslinkers are selected on the basis of their chemical reactivities (i.e., specificity for particular function groups) and other chemical properties that facilitate their use in different specific applications. Choosing among the available crosslinkers can be overwhelming until application needs are considered. By meeting the specific criteria including target functional group, solubility, and cell membrane permeability, the correct crosslinker can be selected for experiments. Reagents are categorized using the general features listed below.

- **Chemical specificity**, including whether the reagent has the same or different reactive groups at either end (e.g., does it have a homo-bifunctional or hetero-bifunctional structure?)
- **Spacer arm length**, including whether the arm is cleavable (e.g., can the linkage be reversed or broken when desired?)
- **Water-solubility and cell membrane permeability** (e.g., can the reagent be expected to permeate into cells and/or crosslink hydrophobic proteins within membranes?)
- **Spontaneously reactive or photo-reactive groups** (e.g., will the reagent react as soon as it is added to a sample or can its reaction be activated at a specific time?)

Recent advances in crosslinking reagents

Crosslinkers can be further classified as either MS-cleavable or MS-non-cleavable. The term cleavable refers to cleavage of the spacer arm within a mass spectrometer. Traditionally, crosslinkers have been the latter. In recent years there has been a push towards novel crosslinkers that cleave during fragmentation within a mass spectrometer.[1,2] This is due, in part, to the challenges associated with analyzing MS data generated using MS-non-cleavable crosslinkers. In a typical MS-based crosslinking experiment using non-cleavable crosslinkers, the data generated consists of both the crosslinked and non-crosslinked peptides, making it very difficult to identify one from the other. There have been number of solutions developed to address the issue such as enrichment of the crosslinked peptides or the use of isotopic labeling.[3-5] In the former, crosslinked peptides are enriched from non-crosslinked peptides, ensuring that the mass spectrometer exclusively samples crosslinked peptides during analysis. Even with sample enrichment, challenges still persist, namely the "n-square problem". In a conventional proteomics experiment, linear peptides are generated upon enzymatic digestion, but a crosslinked peptide generates two linear peptides referred to as α-chain and β-chain linked by a crosslinking reagent upon enzymatic digestion (Figure 3).



$$\Delta m = m_L - m_s = m_{A-L} - m_{A-S} = m_{B-L} - m_{B-S}$$
$$m_A = m_{A-L} - m_L = m_{A-S} - m_S$$
$$m_B = m_{A-B} - m_L = m_{B-S} - m_S$$

Figure 3. Structures of MS cleavable crosslinkers and fragmentation of MS-cleavable crosslinkers within a mass spectrometer.

The MS^2 spectrum that is then generated from the crosslinked peptides consists of fragments from both peptide backbones, with insufficient or unequal fragmentation information for sequencing. When searching crosslinked peptide data using traditional database search approaches, researchers need to consider the possibility of the combination of two peptide sequences plus the reagent itself. This greatly expands the search space as well as the data analysis time (n-square problem). While, there is variety of software developed to address this problem, a better strategy is to employ MS-cleavable crosslinking reagents. In this approach, peptides are crosslinked using reagents that can be cleaved in the gas phase during MS^2 using collision-induced dissociation (CID). This produces two linear peptides in the MS^2 stage that contain parts of the crosslinking reagent differentiating them from the non-crosslinked peptide. These linear peptides can then be targeted for MS^3 acquisition for identification, there by simplifying and speeding up data analysis.

Disuccinimidyl sulfoxide (DSSO) is an ideal reagent for this type of workflow (Figure 3).[1] DSSO is a MS-cleavable crosslinker. It contains an amine-reactive *N*-hydroxysuccinimide (NHS) ester at each end of a 7-carbon spacer arm. The NHS esters react with the primary amine groups on the peptides forming crosslinked peptides. DSSO has similar reactivity to disuccinimidyl suberate (DSS) and bis(sulfosuccinimidyl)suberate (BS^3), but contains a linker that can be cleaved in the gas phase using CID MS^2. The ability to cleave crosslinked peptides during MS^2 enables MS^3 acquisition methods, which can then be used for peptide sequencing using traditional database search engines. The MS cleavage of DSSO generates diagnostic ions during MS^2, which can be used to identify crosslinked peptides from non-crosslinked peptides and in searching using novel database search engines such as XlinkX.[6]

Alternatively, DSBU (Disuccinimidyl Dibutyric Urea) was also introduced with similar thought processes to cleave under CID MS^2 (Figure 3).[2] Like DSSO, DBSU consists of an amine-reactive NHS ester at each end. The spacer arm consists of two aminobutyric acids connected by a central urea moiety. The DBSU reagent can bind to lysines on proteins. Both reagents have number of advantages for crosslinking analysis. They can form number of different crosslinked products. They can form bonds across different proteins or within the same protein, referred to as intra-linked and inter-linked. They can form a number of different crosslinked products and a specific name is given dependent on the type of linking. Crosslinks can form bonds across peptides on different proteins or within two peptides on the same protein, referred to as intra-link and inter-link. Crosslinking can also occur within a single peptide designated as loop-link. The term mono-link is used when one of the NHS ester group is linked to a peptide while the other is hydrolyzed

when no protein/peptide region is in close proximity for the second linkage to occur. The CID fragmentation of these crosslinked peptides can produce diagnostic fragment ions that can be used to ascertain the type of interactions that are occurring.

In general, MS-cleavable crosslinking reagents such as DSSO and DSBU are used as part of a XL-MS workflow to assist in determining partners, domains of protein interactions and three-dimensional structures of proteins. The introduction of MS-cleavable crosslinkers have simplified the whole MS-based workflow by first generating diagnostic ions that can be used as markers to differentiate crosslinked peptides from non-crosslinked peptides, and second as the markers can be targeted for fragmentation to sequence the crosslinked peptides. This has enabled researchers to extract useful information in a timely manner for XI-MS experiments.

Technical advances in mass spectrometry

MS has made it possible for protein-protein interactions to be studied—from simple protein complexes to proteome wide scale experiments, to proteins that were previously inaccessible by traditional techniques. Most importantly, MS has democratized protein interaction studies, making it accessible, inexpensive and high throughput.

There are challenges associated with XL-MS. Specifically, crosslinked peptides tend to be in low abundance relative to un-modified peptides, making detection and characterization by MS problematic. The selection of the mass spectrometer is also crucial for XL-MS as number of issues would hinder most commercial instruments in an XL-MS workflow. Case in point, crosslinked peptides under conventional CID or higher-energy collisional dissociation (HCD) MS^2 fragmentation most commonly available on commercial mass spectrometers generate unequal fragment ions from the peptide backbones. The unequal fragmentation makes it very difficult to sequence the peptides as the peptide backbone would contain insufficient fragmentations for confident identification. The use of multiple fragmentations can overcome the issue with fragmentation of crosslinked peptides, providing improved sequence coverage for both peptide backbones. For example, fragmenting crosslinked peptide with CID and electron transfer dissociation (ETD) in alternate fashion and combining the fragmentation information can improve sequence coverage and increase the total number of crosslinked peptides identified than CID fragmentation alone (Figure 4).[6]

Recently, a novel fragmentation referred to as electron-transfer/higher energy collision dissociation (EThcD) has been introduced.[7,8] In this fragmentation, ETD and HCD are combined in a single spectrum. In EThcD, precursors are fragmented within the linear ion trap using ETD. The precursors, charge reduced precursors and ETD fragment ions are then transferred to the ion-routing multipole (IRM) for HCD fragmentation. The result is an EThcD spectrum containing b-, c-, y- and z- ions, a spectrum that is combination of ETD and HCD fragments. EThcD, similar to ETD, can be acquired in alternating fashion with CID to improve sequence coverage.

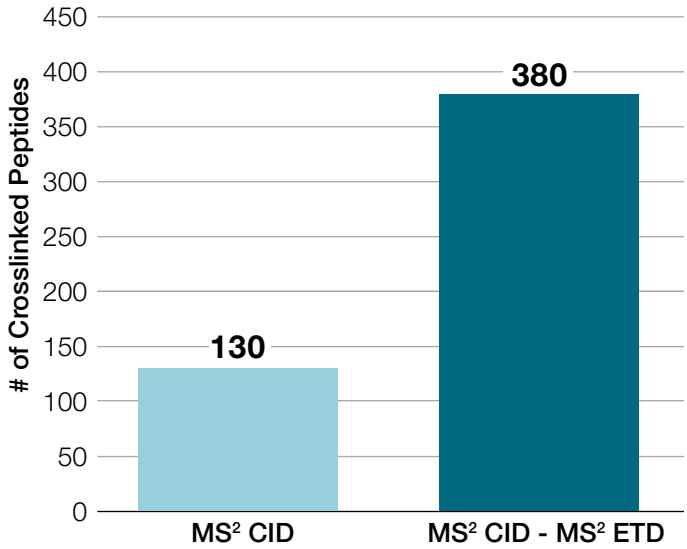


Figure 4. ***Escherichia coli (E.coli)*** **cell lysate crosslinked peptides identified using different acquisition methods.[9]**

Having the ability to perform multiple fragmentations can be beneficial to screen for crosslinked peptides in samples that contain non-crosslinked peptides, particularly when combined with MS^n capability. This allows for screening of crosslinked peptides in an intelligent fashion and acquiring sequence information only when crosslinked peptides are detected. Such an approach can speed up data acquisition and simplify data analysis when combined with MS-cleavable crosslinkers like DSSO.[1] For example, the crosslinking reagent DSSO is shown to generate diagnostic ions in the CID MS^2 spectrum indicative of crosslinked peptides. The unique mass difference generated from the dissociation of DSSO, referred to as the ΔM principle, producing signature peaks within CID MS^2 spectrum, can be used in database searches to filter crosslinked peptides from non-crosslinked peptides.[1]

The signature ions generated from crosslink cleavage are influenced by the type of peptides (identical or different sequences) and the types of interactions (intra-links, inter-links, mono-links or loop-links) that that are linked by DSSO. Such information is beneficial for data analysis as it helps ascertain the type of crosslinking. For example, inter-linked and intra-linked peptides generate four signature ions upon CID fragmentation, if the two crosslinked peptide sequences are different. This is due to the cleavage occurring on both sides of the sulfoxide (C-S bond) within DSSO. Each cleavage site generates two signature ions for a total of four ions with four different masses. However, if DSSO links two identical peptide sequences for interlinked crosslinked peptide then only 2 diagnostic ions are generated.

The diagnostic ions generated in the MS^2 spectrum are masses of the individual peptides present in the crosslinked peptides plus the cleaved DSSO tag. The mass difference produced by the diagnostic ions can be used to trigger MS^n fragmentation for sequencing. Specifically, an HCD MS^3 spectrum can be acquired for these diagnostic fragment ions to sequence the individual crosslinked peptides. Such an approach requires high resolution accurate mass for MS^2 to denote the mass differences as accurately as possible to ensure proper triggering of MS^3. Though CID MS^2 is generated in a nominal mass ion trap mass analyzer, the spectrum detected in the Thermo Scientific™ Orbitrap™ mass analyzer are used to obtain fragments that are of high resolution accurate mass (HRAM).

Different rules apply for loop-linked and mono-linked peptides. For loop-linked peptides, only one diagnostic fragment ion is generated in the MS^2 spectrum as both peptides would still remain linked after cleavage of DSSO. The mass for the fragment ion generated is identical to the MS^1 mass. This peak can then be targeted for MS^3 for sequencing of the crosslinked peptides. The third type of peptides that are generated by DSSO crosslinking is mono-linked modified peptides. This is the case where one end is linked to a peptide while other end of DSSO is not linked to any peptide but rather hydrolyzed. Here the cleavage of the bonds on either side of the sulfoxide generates two fragment ions. These two peptides differ by mass that is the sum of the DSSO fragments still attached to both peptides. Similar to the other two types of crosslinked peptides, these diagnostic ions can then be targeted for MS^3 for sequencing.

XL-MS has benefitted immensely from the latest advancements in mass spectrometers. The availability of multiple fragmentation techniques (CID, HCD, ETD, EThcD), MS^n and the ability to perform these fragmentations at any stage of MS^n enables researchers to increase sequence coverage and thoroughly characterize crosslinked peptides as described above. Newly developed mass spectrometers such as the Thermo Scientific™ Orbitrap Fusion™ Tribrid™ family of instruments have all of these features in a single instrument. thereby enabling researchers to effectively target and characterize crosslinked peptides. Furthermore, the flexibility provided by the unique architecture of new mass spectrometers and the high resolution capabilities present in these instruments enable researchers to develop intelligent acquisition strategies as described above to sequence crosslinked peptides.

Quantitative structural dynamics of protein complexes

In order to understand structural dynamics of protein complexes such as measuring changes in levels of interactions or binding affinities and their place in complex biological systems, quantitative strategies have been developed using stable isotope labeling for XL-MS. However, these approaches suffer from poor co-elution of peptides to limited number of sample comparison. An alternative to stable isotope labeling is the use of isobaric chemical tagging, a popular strategy for relative quantitation in conventional proteomics. In a single analysis, the isobaric labeling approach can be used to identify and quantify relative changes in complex protein samples across multiple experimental conditions. The tags can be used with a wide variety of samples including cells, tissues, and biological fluids. Thermo Scientific™ Tandem Mass Tags (TMT) reagents are isobaric chemical tags consisting of an MS/MS reporter group, balance group, and an amine-reactive group. Amine-reactive groups covalently bind to peptide N-termini and to lysine residues. For crosslinked peptides, TMT efficiently labels the non-crosslinked lysine residue and the free N-terminal primary amines that are produced during enzymatic digestion. After labeling peptides are introduced into the mass spectrometer where each tag fragments during MS^2, producing unique reporter ions. Protein quantitation is accomplished by comparing the intensities of the reporter ions. However, achieving quantitative accuracy is highly dependent on the purity of the precursor ion population selected for MS^2 analysis.

Even with the most rigorous pre-fractionation followed by separation, the issue of co-elution of isobaric species cannot completely be eliminated. The issue of co-isolation and co-fragmentation of interfering ions with ions of interest limit quantitation accuracy and precision. This distortion in TMT ratios influences the final reporter ion population and results in the underreporting of true fold changes and true reporter ion intensities, leading to unpredictable losses of quantitative values. The implementation of synchronous precursor selection (SPS) exclusive to the Tribrid mass spectrometers overcomes these limitations.[10,11] In this approach, the parent ion is selected in MS^1 scan, isolated in the quadrupole and fragment by CID in the ion trap. Upon fragmentation, multiple MS^2 fragment ions are selected and isolated using single trap fill and waveform (synchronous precursor selection). Up to 20 fragments can be isolated simultaneously. Selected MS^2 fragment ions are transferred back into the IRM and HCD fragmentation is performed. MS^3 fragments are detected in the Orbitrap analyzer for the most accurate MS^3 based TMT quantitation. Using SPS dramatically increases the signal intensity and improves the ratio accuracy (due to counting statistics) and at the same time dramatically boosts sensitivity increasing the total number of peptides quantified significantly. Combining the isobaric chemical tagging approach with XL-MS alleviates the problems that has hindered quantitative crosslinking-MS analysis. Since quantitation is performed at the MS^n level, co-elution issues that plague stable isotopic approaches are no longer an issue. Additionally, samples can be multiplexed, up to eleven samples using commercially available reagents.

Recent advances in quantitative XL-MS

A novel workflow has been recently developed for XL-MS that taps into the multiplexing capability offered by TMT for quantitation of crosslinked peptide.[12] This workflow referred to as quantitation of multiplexed, isobaric-labeled cross (X)-linked peptides or QMIX combines targeted MS^3 triggering of diagnostics ions in MS^2 as previously described for identification with SPS MS^3 for quantitation (Figure 5). This strategy enables quantitation up to 11 different conditions in a single analysis and enables protein-protein interaction studies in a proteome wide scale.

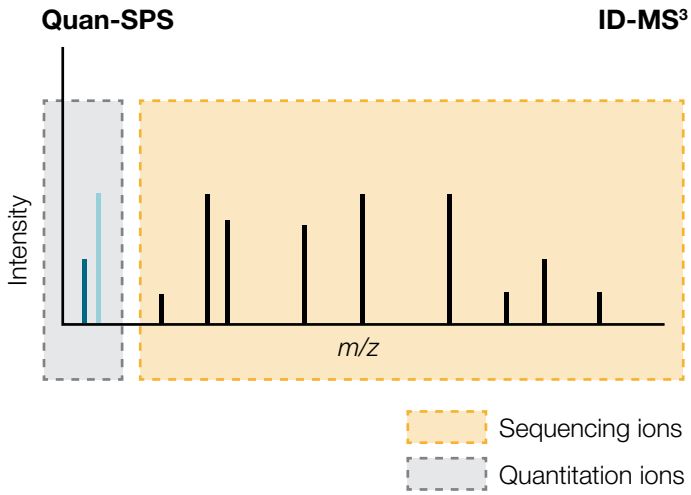


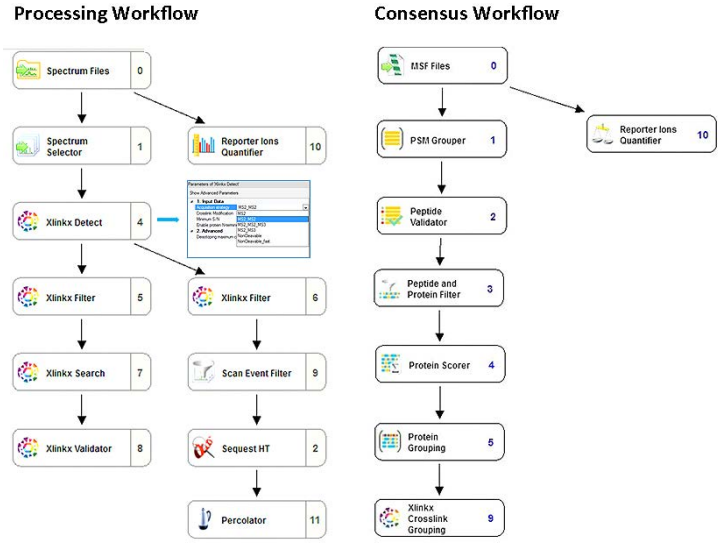
Figure 5. Quantitation of multiplexed, isobaric-labeled cross (X)-linked peptides or QMIX workflow.

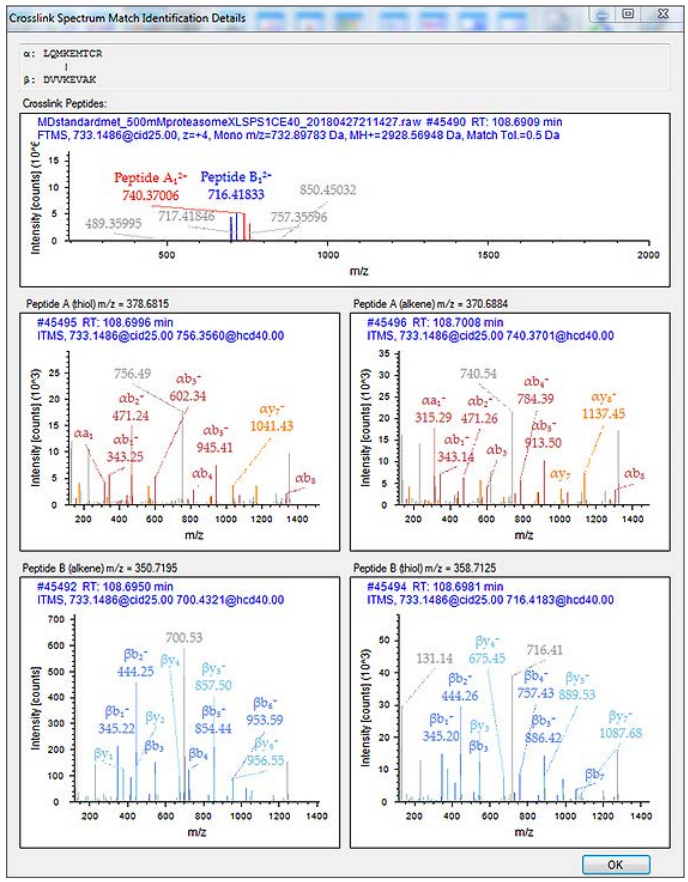
Solving the bottleneck of data analysis

As described earlier data analysis in XL-MS workflows is compounded by the n^2 problem and most commercial software are not designed to handle XL-MS data. Researchers are left to design in-house software to process XL-MS data. Unfortunately, such software are not available to the community at large. Issues such as ease of use, support by developer and stability of the software tend to plague these types of development. The ideal XL-MS software should be an extension of current proteomics software as the data that is generated have a lot in common with traditional bottom-up proteomics data.

Due to this reason, a new algorithm called XlinkX has been developed to address this issue.[13] XlinkX takes advantage of the unique signature ions that are produced during the cleavage of the MS-cleavable crosslinker during MS^2 acquisition. These peaks provide the masses of the crosslinked peptides thus preventing the algorithm from trying all combinations of the peptides of a proteome to find the masses of the two peptides that form the crosslinker. This converts the n^2 growth of the search space with the number of peptides searched into a problem with the standard linear growth of a sequence database search algorithm. Thus, reducing analysis time significantly. The XlinkX software supports multiple crosslinker reagents as well as different fragmentations (CID, HCD, ETD and EThcD). A XlinkX node has been developed for the widely used Thermo Scientific™ Proteome Discoverer™ software (Figure 6). Processing the data within within Proteome Discoverer enables researchers to export XL-MS results for validation in xiNET, which is a visualization tool for exploring XL-MS results.

Despite the advances in reagents and mass spectrometry, data analysis has hindered XL-MS and prevented it from becoming a widely used workflow. The development of XlinkX has addressed this workflow and makes it accessible to the community at large. Not only does XlinkX speeds up data analysis, but more importantly it has been incorporated into a widely used and familiar proteomics software such as Proteome Discoverer software.





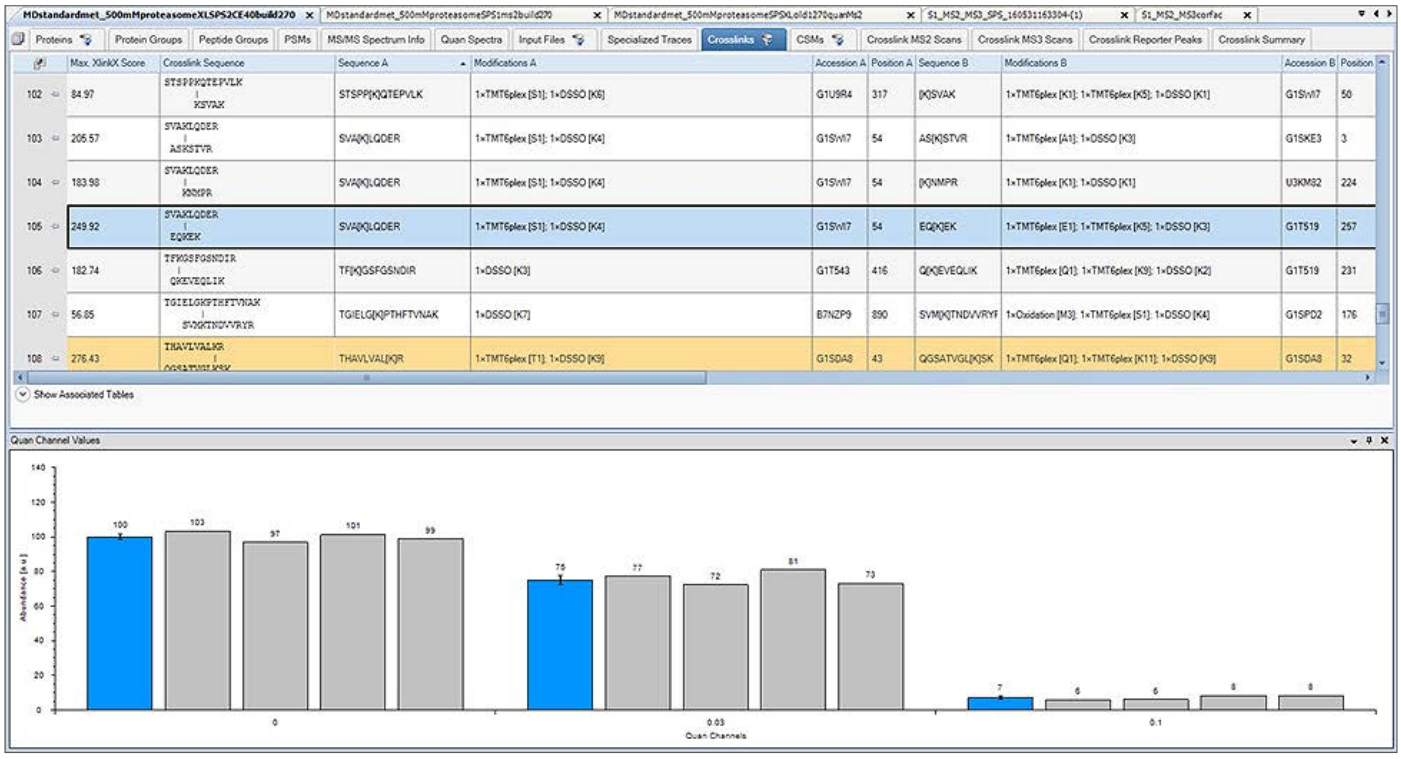


Figure 6. A typical processing and consensus workflow to identify crosslinked peptides using the XlinkX nodes is shown along with proteins, quantitation results tab and spectra annotation.

thermo scientific

Conclusion

XL-MS enables analysis of protein-protein interactions to better understand how proteins affect biological processes such as signaling cascades, gene upregulation, and energy (ATP) production It is via such interactions that biological processes commence, conclude, and change. Therefore, the study of protein interactions using XL-MS is instrumental in understanding both healthy and disease states. Primary advantage of XL-MS is that it is accessible to proteomics laboratories without purchase of additional equipment. It also complementary to alternate techniques such as cryo-EM, X-ray crystallography, NMR and MS techniques such as HDX, native MS. XL-MS enables the examination of interactions close to the physiological state of an organism, generating interaction information that is biologically relevant.

The development of MS-cleavable crosslinking reagents such as DSSO and DSBU have spearheaded a revolution in XL-MS. By taking advantage of how a mass spectrometer works the reagents have improved the approach to data acquisition and simplified data analysis. The reagents crosslink interacting regions of a protein or protein complex and during MS^2 fragmentation the reagents cleave to produce two linear peptides that can be easily sequenced and identified. Mass spectrometers with flexible architecture and novel fragmentations such as ETD/EThcD have taken the workflow to the next level. Enabling intelligent acquisition strategies that provide improved identification and sequence coverage for crosslinked peptides. And finally commercially available software such as XlinkX have made the workflow accessible to the community providing a search engine that can identify, sequence in a timely manner.

References

1. Kao, A.; Chiu, C. L.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L., Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol Cell Proteomics* **2011**, *10* (1), M110 002212.
2. Muller, M. Q.; Dreiocker, F.; Ihling, C. H.; Schafer, M.; Sinz, A., Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS. *Anal Chem* **2010**, *82* (16), 6958–68.
3. Yu, C.; Kandur, W.; Kao, A.; Rychnovsky, S.; Huang, L., Developing new isotope-coded mass spectrometry-cleavable cross-linkers for elucidating protein structures. *Anal Chem* **2014**, *86* (4), 2099–106.
4. Yu, C.; Mao, H.; Novitsky, E. J.; Tang, X.; Rychnovsky, S. D.; Zheng, N.; Huang, L., Gln40 deamidation blocks structural reconfiguration and activation of SCF ubiquitin ligase complex by Nedd8. *Nat Commun* **2015**, *6*, 10053.
5. Boutilier, J. M.; Warden, H.; Doucette, A. A.; Wentzell, P. D., Chromatographic behaviour of peptides following dimethylation with H2/D2-formaldehyde: implications for comparative proteomics. *J Chromatogr B Analyt Technol Biomed Life Sci* **2012**, *908*, 59–66.
6. Liu, F.; Rijkers, D. T.; Post, H.; Heck, A. J., Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat Methods* **2015**, *12* (12), 1179–84.
7. Frese, C. K.; Altelaar, A. F.; van den Toorn, H.; Nolting, D.; Griep-Raming, J.; Heck, A. J.; Mohammed, S., Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal Chem* **2012**, *84* (22), 9668–73.
8. Frese, C. K.; Zhou, H.; Taus, T.; Altelaar, A. F.; Mechtler, K.; Heck, A. J.; Mohammed, S., Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (EThcD). *J Proteome Res* **2013**, *12* (3), 1520–5.
9. Bomgarden, R.; Raja, Erum.; Etienne, C.; Liu, F.; Heck, A.; Mueller, M; Viner, R., Optimization of crosslinked peptide analysis on an Orbitrap Fusion Lumos mass spectrometer, ASMS 2016 Poster.
10. Ting, L.; Rad, R.; Gygi, S. P.; Haas, W., MS^3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods* **2011**, *8* (11), 937–40.
11. McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, I.; Grothe, R. A.; Blethrow, J. D.; Gygi, S. P., Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal Chem* **2012**, *84* (17), 7469–78.
12. Yu, C.; Huszagh, A.; Viner, R.; Novitsky, E. J.; Rychnovsky, S. D.; Huang, L., Developing a Multiplexed Quantitative Cross-Linking Mass Spectrometry Platform for Comparative Structural Analysis of Protein Complexes. *Anal Chem* **2016**, *88* (20), 10301–8.
13. Liu, F.; Lössl, P.; Scheltema, R.; Viner, R.; Heck, A.J.R., Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nature Com* **2017**, *8*, 15473.

Find out more at **thermofisher.com/ms-structure**

ThermoFisher SCIENTIFIC

© 2018 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is presented as an example of the capabilities of Thermo Fisher Scientific products. It is not intended to encourage use of these products in any manners that might infringe the intellectual property rights of others. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representatives for details.

WP65343-EN 0918M