Published in final edited form as:

J Proteome Res. 2019 April 05; 18(4): 1916–1925. doi:10.1021/acs.jproteome.9b00114.

Evaluation of Different Stationary Phases in the Separation of Inter-crosslinked Peptides

Zixiang Fang¹, Yehia Z. Baghdady¹, Kevin A. Schug¹, and Saiful M. Chowdhury^{1,*}

¹Department of Chemistry & Biochemistry, The University of Texas at Arlington, Arlington, TX, USA

Abstract

Chemical crosslinking coupled with mass spectrometry (MS) is becoming a routinely and widely used technique for depicting and constructing protein structures and protein interaction networks. One major challenge for crosslinking/MS is the determination of informative low abundant intercrosslinked products, generated within a sample of high complexity. A C18 stationary phase is the conventional means for reverse-phase (RP) separation of inter-crosslinked peptides. Various RP stationary phases, which provide different selectivities and retentions, have been developed as alternatives to C18 stationary phases. In this study, two phenyl-based columns, biphenyl and fluorophenyl, were investigated and compared with a C18 phase for separating BS³ (bis(sulfosuccinimidyl)suberate) crosslinked bovine serum albumin (BSA) and myoglobin by bottom-up proteomics. Fractions from the 3 columns were collected and analyzed in a linear ion trap (LIT) mass spectrometer for improving detection of low abundant inter-crosslinked peptides. Among these three columns, the fluorophenyl column provides additional ion exchange interaction and exhibits unique retention in separating the crosslinked peptides. The fractionated data was analyzed in pLink, showing fluorophenyl column consistently obtained more inter-crosslinked peptide identifications than both C18 and biphenyl columns. For BSA crosslinked sample, the identified inter-crosslinked peptide number of fluorophenyl to C18 column are 136 to 102 in “low confident” result and 11 to 6 in “high confident” result. Fluorophenyl column could potentially be a better alternative for targeting the low stoichiometric inter-crosslinked peptides.

Graphical Abstract

*Corresponding authors Saiful M Chowdhury, PhD Associate Professor Chemistry and Biochemistry University of Texas at Arlington Box 19065 700 Planetarium Place, Room 130 Arlington, TX 76019 Tel: 817-272-5439 Lab:817-272-2071 schowd@uta.edu, <http://www.uta.edu/chemistry/faculty/directory/Chowdhury.php>, <https://www.uta.edu/profiles/saiful-chowdhury>.

The authors declare no competing financial interests.

Associated Contents

Supporting Information:

These Supporting Informations are available free of charge via the Internet at <http://pubs.acs.org>.

Keywords

Chemical crosslinking; Mass spectrometry; Phenyl-based stationary phases; Fluorophenyl column; Crosslinked peptides; Peptides separation and enrichment

Introduction

Chemical crosslinking is on its way to becoming a routine and straightforward technique for exploring protein interactions and protein structures. In a crosslinking study, a crosslinker is used to chemically react, derivatize, and covalently link the side chains of specific amino acids, such as lysine, cysteine or other charged residues, between interacting proteins under physiological conditions.^{1–3} An efficient crosslinking reaction can capture and stabilize macromolecules when two reactive side chains of interacting proteins are within the reacting range of the crosslinker.⁴ By customizing the reactive groups and the span of a crosslinker, the accessibility of certain amino acids in proteins and the distance of protein interactions can be deduced and utilized for determining the low-resolution structures of proteins and their interacting partners.^{4–7} The simplicity of experimental setup and the capability to capture weak/transient interactions render crosslinking a great complementary approach to traditional high-resolution techniques, such as X-ray crystallography or NMR.^{7–9} Coupled with sensitive mass-spectrometric detection, chemical crosslinking is developed to be a very powerful tool for identifying protein interactions and unravelling the structures of proteins or protein complexes.^{10–16}

Despite many advantages with respect to deducing protein structures, chemical crosslinking still faces certain challenges and limitations. The major drawback is the low abundance of crosslinked proteins. Chemical crosslinking studies are generally performed in the fashion of “bottom-up” proteomics, which begins with the crosslinking reaction of proteins in biological context, followed by enzymatic digestion of proteins into peptides and further analysis of these peptides by LC-MS/MS.^{1, 17} These crosslinked peptides generated by reacting with crosslinkers are not equally informative.^{6, 17} Three main types of crosslinked peptides can be formed after crosslinking and enzymatic digestion of proteins:¹⁸ (1) Intra-crosslinked peptides; (2) dead-end crosslinked peptides; and (3) inter-crosslinked peptides. An intra-crosslink or “loop-link” is the crosslinking of two reactive groups within the same

peptide sequence; a deadend crosslink or “mono-link” is the crosslinking of one reactive group, while the other reactive end is deactivated; an inter-crosslink is the crosslinking of two functional groups of two different peptide sequences. The information on protein conformation and interaction is predominately acquired from the inter-crosslinked products. However, inter-crosslinked peptides usually account for the smallest portion while dead-end counterparts prevail following execution of the crosslinking reaction.¹⁷ Enzymatic dissection generates enormous numbers of unmodified peptides which further lowers the relative abundance of inter-crosslinked peptides. On the other hand, the interpretation and analysis of crosslinking MS data is considered much more timeconsuming and computationally-intensive than traditional proteomic analysis.⁵ The difficulty in interpreting MS data associated with crosslinking experiments is due to the combinatorial increase from a theoretical non-crosslinked peptide number of n , to a theoretical crosslinked peptides number of n^2 .¹⁷ This requires more computational capacity in hardware and complicated searching algorithms in software for crosslinking analysis, especially for samples with high biological complexity. Many software tools addressing the data analysis of cleavable or noncleavable crosslinking/MS have been developed, such as MetaMorpheus¹⁹, X-linkX¹³, Xlink Identifier²⁰, SIM-XL²¹, Kojak,²² Stavrox,²³ and pLink²⁴. However, without efficient enrichment of inter-crosslinked peptides prior to mass spectrometric analysis, the processing of MS data to confidently identify inter-crosslinked peptides remains still a challenging goal.

Strategies had been designed for reducing complexity in crosslinked samples and enriching inter-crosslinked peptides. Trifunctional crosslinkers, which incorporate an affinity tag like a biotin moiety, can be specifically targeted and selectively enriched by immobilized avidin.^{2, 25, 26} Dual cleavable crosslinker DUCCT was developed to pinpoint the inter-crosslinked peptides by simultaneous CID and ETD fragmentations.⁵ After the crosslinking reaction, noncrosslinked peptides can be removed before the recovery of crosslinked peptides from avidin. However, this approach cannot differentiate dead-end crosslinked peptides from intercrosslinked peptides, and the non-specific binding of unmodified peptides may occur. Meanwhile, the biotin group can also be fragmented in tandem MS. This further complicates MS analysis and increases false discovery rate (FDR). Structurally, inter-crosslinked peptides can be considered as the sum of two individual linear peptides fixated by a crosslinker. These products are structurally having higher molecular mass and longer average sequence than dead-end crosslinked and unmodified peptides. Thus, size exclusion chromatography (SEC) is applicable to enrich and separate inter-crosslinked peptides from their non-crosslinked counterparts.^{11, 27} Moreover, inter-crosslinked peptides bear higher charges in acidic environment, because they possess more chargeable groups within the combined peptide sequence, relative to noncrosslinked species. Thus, strong cation exchange (SCX) chromatography is also feasible to separate highly charged inter-crosslinked peptides from other types of peptides.^{8, 28, 29} Fritzsche et al. compared the number of inter-crosslinked peptides of BS³ crosslinked BSA with or without SCX enrichment and observed almost an 8-fold increase in inter-crosslinked-peptide identifications after enrichment.²⁸ More recently, two-dimensional (2D) SCX approaches were developed to further assist in the separation of inter-crosslinked peptides. Buncherd et al. explored a chemical-cleavable crosslinker and utilized 2D diagonal SCX chromatography for enriching inter-crosslinked peptides.²⁹ The determination of inter-crosslinked peptides was achieved by mapping the

relation of intact crosslinked peptide before and after chemical cleavage. Another 2D SCX approach was reported by Tinnefeld et al.⁸ In that study, LysC digestion of crosslinked protein was performed to generate long, highly charged peptides, which were preferentially retained in the first dimension SCX separation. Upon the first SCX fractionation, the lengthy crosslinked peptides were subjected to tryptic digestion and followed by a second SCX separation and fraction collection to achieve efficient enrichment.

Alkyl-based stationary phases are the most widely utilized bonded phase in RP highperformance liquid chromatography (HPLC) for the separation of crosslinked peptides. All the aforementioned enrichment methods from the literature inevitably included C18 separation during the separation of crosslinked peptides, regardless of the sample complexity. In a C18 column the crosslinked peptides are separated based on their relative hydrophobicity. In general, the hydrophobicity of a peptide increases as the number of amino acids in the peptide sequence increases. However, the increase of basic ionizable functional groups such as histidine, lysine, and arginine in longer peptides will increase the number of accessible charge states and yield more interactions with charged stationary phases under acidic condition. Another major difference between the inter-crosslinked peptides and others is that the geometric arrangements of inter-crosslinked peptides are changed when covalently linked. This conformational distinction of inter-crosslinked peptides has the potential to be exploited for targeted separation using stationary phases with shape selectivity.

Many alternative stationary phases to C18 have been developed. These stationary phases have been compared and shown variations in selectivity and retention for multiple classes of analytes,³⁰ but they have not been fully explored for separating complex mixture of macromolecules, particularly crosslinked macromolecules. Here, we assessed C18, biphenyl, and fluorophenyl columns for their ability in selectively separating and enriching inter-crosslinked peptides of BSA crosslinked by BS³ crosslinker. We evaluated and compared the performance of three commercially-available column chemistries according to chromatographic retention and number of identified inter-crosslinked peptides by the crosslinking software “pLink1”.²⁴ This work aims to provide insights in improving separation and enrichment of the low abundant intercrosslinked peptides by alternative stationary phases.

Materials and Methods

Materials and Reagents

Bovine serum albumin (BSA), myoglobin of equine heart, Tris-HCl, ammonium bicarbonate (ABC), and formic acid (FA) were purchased from Sigma-Aldrich (St. Louis, MO). Methanol and acetonitrile were obtained from VWR (Radnor, PA). BS³ crosslinker was obtained from Thermo Scientific (Rockford, IL). 3K MWCO protein concentrators (Pierce, IL) were used to concentrate the protein after the crosslinking reaction. Dithiothreitol (Biorad, CA), iodoacetamide (Sigma-Aldrich, MO), and trypsin (Promega, Madison, WI) were obtained. MilliQ filtrated water (resistivity > 18 MΩ) by Aries Filterworks (West Berlin, NJ) filtration system was used to prepare all aqueous solutions in this study.

Crosslinking of BSA

BS³ crosslinker was prepared by dissolving to 50 mM solution in water. BSA was reacted with BS³ crosslinker in 1:50 molar ratio in pH 7.3 phosphate buffer saline (PBS). Different concentrations of crosslinked samples were achieved by increasing the amount of BSA used for the crosslinking reaction. The reaction was performed under room temperature for 30 min in a constantly mixing thermomixer. The crosslinking reaction was quenched by adding Tris-HCl buffer (pH 8.8) to a final concentration of 40 mM. The sample was transferred to a 3K MWCO protein concentrator to remove excess crosslinker by washing with 50 mM ABC solution under 15,000 G centrifugation. Then, the crosslinked BSA sample was subjected to in-solution digestion.

In-solution Digestion

The crosslinked BSA protein was first reduced with 10 mM dithiothreitol solution and then alkylated with 55 mM iodoacetamide in dark. The digestion was carried out by incubating the sample with modified trypsin in 1:100 (w/w) trypsin/protein ratio at 37°C overnight. After that, the enzymatic digestion was stopped by adding FA. The digested sample was dried by a speed vacuum concentrator and resuspended in 0.1% FA solution for LC separation and mass spectrometric analysis.

LC-MS/MS analysis

Analysis of the crosslinked BSA samples using C18, biphenyl, and fluorophenyl columns was performed on a Prominence HPLC system coupled to an electrospray ionization (ESI) ion trap (IT) time-of-flight (TOF) mass spectrometer (Shimadzu, Japan). The chromatographic separation in the HPLC-ESI-IT-TOF was achieved using three Restek (Bellefonte, PA, USA) columns: Raptor™ C18 column (100 × 2.1 mm, 2.7 μm), Raptor™ Biphenyl column (100 × 2.1 mm, 2.7 μm) and Force™ Fluorophenyl column (100 × 2.1mm, 3μm) in a 60-min gradient (Flow rate: 0.2 mL/min; Gradient: 0–2 min 5% B, 2–40 min 5–50% B, 40–40.01 min 50–100% B, 40.0150 min 100% B, 50–50.01 min 100–5% B, 50.01–60 min 5% B; Mobile phase A: 0.1% FA in water; Mobile phase B: 0.1% FA in 95% acetonitrile, 5% water). MS scan mass range for MS and MS/MS was set from m/z 200 to 2000. Data-dependent acquisition was performed for the fragmentation of peptides. Five MS events (one MS scan and four MS/MS scans) were performed. Event 1 was full MS scan. Event 2 to 5 (4 MS/MS scan events) selected the four most abundant ions of each event 1 scan for further fragmentation (charge state 2; isolated width of 3 Da; precursor ion threshold: 20000). Collision induced dissociation (CID) energy was set to 60% and collision gas was 50%. Dynamic exclusion was set to 10 sec.

Improved detection of crosslinked samples was carried out by collecting fractions from three columns. Each fraction was collected using the previous HPLC system. The fractions were analyzed in a Dionex UltiMate 3000 nano-UHPLC system coupled with nano-ESI-linear ion trap (LIT) Thermo Velos Pro mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). The separation in the nano-LC-ESI-LIT was achieved by Acclaim™ PepMap™ C18 column (150 mm × 75 μm, 3μm) in a 70 min gradient separation (Flow rate: 0.300 μL/min; 70 min Gradient: 0–3 min 4.0%B, 3–50min 4.0–50.0% B, 50–50.1 min 50–100% B, 50.1–55 min 100% B, 55–55.1 min 100–4% B, 55.1–70 min 4% B; Mobile phase A: 0.1% FA in

water; Mobile phase B: 0.1% FA in 95% acetonitrile, 5% water). MS data was obtained from 200 to 2000 m/z mass range. Data-dependent MS/MS spectra were collected from the five most abundant precursor ions upon fragmentation (charge state 2; isolated width of 2 Da; min. signal required: 500) using CID activation with 35.0% normalized collision energy, activation Q of 0.25, and activation time of 30 ms.

Identification of Inter-Crosslinked Peptides by pLink

The original MS data files were converted to mgf files by Mascot (for Shimadzu lcd files) or Proteome Discoverer (for Thermo raw files). Identification of inter-crosslinked peptides was performed using pLink1 crosslinking software.²⁴ The FASTA database of BSA sequence was acquired from Uniprot. The parameters for pLink search were as follow: sample1.spectra. instrument=CID; sample1.spectra. format = mgf; mod. fixed. 1=Carbamidomethyl[C]; mod. variable.1=Oxidation[M]; linker.name1=BS3; noninterexport=true; value_max=0.1. Mass tolerances of precursor and product ion searching are either 100 ppm or 1000 ppm depending on the type of mass spectrometer used for detection. The other parameters were set as default. Results of the identification were annotated with b and y ions of peptide sequence and visualized by pLabel software, a component of the pLink software.

Results and Discussions

Evaluation the Separation of Crosslinked BSA in Three Columns

Currently, alkyl C18-bonded stationary phases are still the standard stationary phase in RPLC for separating crosslinked samples. These generally show consistent performance in separating hydrophobic compounds, but they are found not to be as efficient for certain compounds, specifically those that are more hydrophilic in nature.³¹ Many alternative stationary phases have been developed to improve retention and selectivity of targeted analytes.³² These generally rely on the introduction of addition functional units into the stationary phase that can then interact with analytes through a myriad of intermolecular forces. Phenyl-based stationary phases were developed to provide π - π interaction mechanisms, as well as to generate separative driving forces from polarity and shape selectivity.^{33, 34} Biphenyl stationary phases have been shown to exhibit increased retention and enhanced selectivity towards polarizable or hydrophilic aromatic compounds. Fluorophenyl stationary phases have been shown to provide increased retention for basic hydrophilic compounds. Additional interactive properties of the fluorophenyl column were previously investigated in the applications of off-line fractionations on Hela cells³⁵, orthogonality in 2D-LC separations³⁶ and separation of metabolites in urine samples³⁷. Fluorine atoms incorporated in the fluorophenyl stationary phase act as electron withdrawing groups and induce enhanced interactions between the aromatic group and positively charged compounds. More importantly, the shape selectivity from the phenyl columns could be beneficial in separating the “H”-shaped inter-crosslinked peptides from unmodified linear peptides or dead-end crosslinked peptides. Thus, these two columns were selected to evaluate their potential for improved separation of cross-linked peptides, relative to a traditional C18 phase. To the best of our knowledge, no prior literature reported on the separation of crosslinked peptides by biphenyl and fluorophenyl phases, though the

structurally unique inter-crosslinked peptides appears appealing and promising in separation by alternative stationary phases other than C18.

Different concentrations of BSA crosslinked samples were initially investigated in HPLC-ESI-IT-TOF mass spectrometer by these three columns. Three concentrations (0.6 $\mu\text{g}/\mu\text{L}$, 1.7 $\mu\text{g}/\mu\text{L}$ and 5.0 $\mu\text{g}/\mu\text{L}$, measured by Thermo Scientific Nanodrop 2000c instrument) of the crosslinked BSA samples were prepared in triplicate. Each replicate was injected three times for each column and performed MS acquisition as described in the experimental section. The numbers of inter-crosslinked peptides identified are summarized in Table 1 and the list is shown in Supplementary Table S1. The 1.7 $\mu\text{g}/\mu\text{L}$ samples yielded the most identifications of intercrosslinked peptides among all three columns and were further investigated to compare the performance of the three columns for separating inter-crosslinked peptides (total 9 runs for each column).

Total ion current (TIC) chromatograms obtained from 1.7 $\mu\text{g}/\mu\text{L}$ samples by three columns are shown in Figure 1. Chromatograms of 3 biological replicates from the same column are consistent in peak distributions, confirming the reproducibility of the crosslinking method and column performances. When comparing the separative performances of each column based on chromatograms, most peaks eluted within the first 30 min for the C18 and biphenyl columns, while the peaks were more broadly distributed and retained until 35 min on the fluorophenyl phase, using the same mobile phase conditions. The fluorophenyl phase showed unique retention and separative pattern, enabling elongated retention for certain peptides in the crosslinked BSA sample over a longer time range.

Table 2 summarizes six inter-crosslinked peptides that were mutually identified from the pLink analysis by all three columns. Comparing the retention times of these inter-crosslinked peptides, they were eluted around the same retention times in C18 and biphenyl columns. These inter-crosslinked peptides were slightly less retained on the biphenyl column compared to the conventional C18 column, with the maximum difference of 0.8 min. In the separation on the fluorophenyl column, the peptides were significantly more retained compared with the C18 and biphenyl phases. The retentions of the crosslinked peptides were increased by between 2.5 min to more than 4 min in retention time in the fluorinated phase. Similar retention shifts of noncrosslinked peptides (unmodified and dead-end crosslinked peptides) were also observed (Supplementary Table S2). From the differences in retention times we find that the intercrosslinked peptides are retained in C18 and biphenyl columns similarly, whereas they exhibited more interactions in fluorophenyl column. This could be related to the charge-based interactions between the negative dipole in fluorine and positively charged inter-crosslinked peptides. The fact that inter-crosslinked peptides are more interactive with the fluorophenyl phase could be beneficial for improving their separation and identification, relative to the use of a traditional C18 column.

From pLink software, both search results of inter-crosslinked and non-crosslinked peptides (unmodified and dead-end crosslinked peptides) were reported and compared in evaluating the column performance. In the data-dependent mode of MS analysis, the more abundant precursor ions have higher chances of being selected for the fragmentation. Thus, the better a certain type of peptides is separated within the column, the more identifications can be

achieved and reported. We expect the column with the most inter-crosslinked peptide identifications having the best separation of inter-crosslinked peptides. In contrast, the naturally more abundant non-crosslinked peptides are undesired because they act as “matrix” in the sample and co-elute with more informative inter-crosslinked peptides, competing for fragmentation. The less identification of non-crosslinked peptides shows that the “matrix” is more depleted, and the crosslinked peptides could be relatively preferred. When we investigated the number of non-crosslinked peptides identified in total 9 runs from each of three columns, the biphenyl column had 48.0 ± 2.8 , the C18 column achieved 52.0 ± 5.0 while fluorophenyl column only had 40.1 ± 3.3 in non-crosslinked peptide identifications (Data not shown). The number of non-crosslinked peptides identified in the fluorophenyl column is statistically less than the other two columns which helps to reduce the detection and fragmentation of undesired non-crosslinked peptides.

In terms of inter-crosslinked results, the number of identifications for the biphenyl, C18 and fluorophenyl column in three replicates from 9 runs were calculated to be 5.1 ± 0.9 , 5.4 ± 1.1 and 4.4 ± 1.2 respectively. Considering the deviation in relation to number of identified peptides, there is no significant advantage for any column in obtaining more inter-crosslinked identifications, considering the low sensitivity of MS instrument. Therefore, this result is not conclusive for determining the column performance. A nano-LC-ESI system would be preferred for separating and detecting the low abundant crosslinked samples with high dynamic range. However, there are no commercially available capillary biphenyl or fluorophenyl columns compatible with nano-LC system available to date. We are in the process of constructing the column with fluorophenyl phase.

Fraction Collections and Analysis by Nano-UHPLC-ESI-LIT

According to the previous analysis of inter-crosslinked results, three columns performed almost identical in identifying inter-crosslinked peptides under regular HPLC-ESI-IT-TOF mass spectrometer, regardless of the retention differences on the different phases. The evaluation of columns was mainly hampered by low sensitivity considering the numbers of inter-crosslinked peptides identified were less than previous reported.^{8, 38, 39} Therefore, system with higher sensitivity is necessary for more appropriate evaluation of performances of three columns. We improved the sensitivity of detection by collecting the fractions from each column and analyzing the fractions in a more sensitive nano-UHPLC-ESI-LIT instrument. During the sample fractionation, the MS setting in LC-ESI-IT-TOF using the three columns were maintained the same. Based on the chromatograms (Figure 1), fractions were collected on the retention times containing the time span when the inter-crosslinked peptides were eluted from the columns. For C18 and biphenyl columns, most peaks were eluted within the first 30 min. Thus, the fractions were collected from 0–31 min of the LC runs for C18 and biphenyl columns. To reduce analysis time of fractions, the collection of fractions was performed by combining first 10 min eluent as one fraction and every 3-minute interval for retention time 10–31 min. A total of 8 fractions (0–10 min; 10–13 min; 13–16 min; 16–19 min; 19–22 min; 22–25 min; 25–28 min; 28–31 min) were collected from separations performed on the C18 and biphenyl columns. Fraction collection for the fluorophenyl column was performed in the same fashion, except that the collection time was from 0–37 min in accordance with the extended retention times. A total of 10 fractions (0–

10 min; 10–13 min; 13–16 min; 16–19 min; 19–22 min; 22–25 min; 25–28 min; 28–31 min; 31–34 min; 34–37 min) were collected for the fluorophenyl column. These fractions were dried, reconstituted, and analyzed in the nano-UHPLC-ESI-LIT system as described in the method.

The increase of sensitivity was reflected in the increasing matching of fragments for confident identification of inter-crosslinked peptides (Figure 2). After nano-LCMS analysis and pLink search, the number of inter-crosslinked peptides identified from each fraction of each column was determined. Under the same searching parameters, the number of inter-crosslinked peptides identified had increased significantly after collecting fractions. By comparing the number of unique peptides, the fractions from the C18 column resulted in the identification of 59 unique inter-crosslinked peptides, the fractions from the biphenyl column returned 82 unique inter-crosslinked peptides, and fractions from the fluorophenyl column resulted in 102 unique inter-crosslinked peptides (Figure 3a).⁴⁰ Both the biphenyl and fluorophenyl phases returned more inter-crosslinked peptides than the C18 column.

However, very low percentage of the total inter-crosslinked peptides is mutually identified by all three columns as seen in the Venn diagram. We suspect the reason could be the relatively low resolution in the ion trap type mass analyzer and if using the same searching parameters as TOF instrument, some inter-crosslinked peptide candidates could be filtered by the limited tolerance of precursor and fragment ions. Thus, we re-analyze the same data with a higher tolerance of 1000 ppm for precursor and fragment ion matching. In terms of the overall identified crosslinked peptides each column had increased from 204 to 233, and more importantly, the mutually identified inter-crosslinked peptides increase to 12.4% in percentage from 4.4% in the first search (Figure 3b). Fluorophenyl column still yields the greatest number of inter-crosslinked identifications, identified 136 compared to 102 from C18 and 101 from biphenyl column, outperforming the other two columns by the increase of more than 30% in inter-crosslinked peptides identified.

During each analysis by the pLink software, multiple result files are generated representing different degree of filtering by the default 5% FDR. One result of the search is generated after filtering each spectra file independently, from which the reported intercrosslinked number is shown in Figure 3a and 3b. There is also a re-filtered result file combining all the independent filtered result files and generate a combined result file which provides more confident identifications. When comparing the combined results from each column, the identification numbers are largely reduced. Only 18 inter-crosslinked peptides in total are reported in the “high confident” results from all columns studied. These identified peptides are listed in Table 3. Regardless of the stricter filter, the greatest number of intercrosslinked peptides came from the fluorophenyl column, reporting 11 out of 18 totals identified against 6 from C18 column and 9 from biphenyl column. Furthermore, for validating the plausibility in forming the reported inter-crosslinked peptides and evaluating the columns in identifying “true” crosslinked peptides, distance between two reacted lysines of each crosslinked peptide was calculated in reference to the crystal structure of BSA from protein data bank(entry: 3V03, Supplementary Figure S1). The maximum distance between two α -carbons of BS³ crosslinked lysine is reported to be 27.4 Å, taken into account the span of BS³ crosslinker, two lysine side chains and side chain flexibility.²⁸ A total of 7 of the

fluorophenyl identified peptides are within the calculated limit of BS³ crosslinker with 4 for either of the other two columns (Table 3). This confirms that fluorophenyl column is able to achieve better separation and identification of the “true” crosslinked peptides.

In addition, myoglobin from equine heart (around 17 kDa molecular weight) was also studied by the fractionation method as in BSA fractionation. There is no remarkable improvement in biphenyl column, thus C18 column and fluorophenyl columns are compared by the “high confident” results from pLink software. Fluorophenyl continues to outperform the C18 column by identifying 5 inter-crosslinked peptides, from which 4 out of 5 within the maximum distance of BS³ crosslinking (Supplementary Figure S2). For C18 column, only 1 intercrosslinked peptide which crosslinked in the same Lys147 position (possibly from dimer) was reported (Supplementary Table S3). The superiority of fluorophenyl column to C18 column in crosslinking separation is further confirmed.

Nano-LC-MS improved the overall sensitivity to address low abundant crosslinked peptides, which assisted in appraising the three columns in separating and identifying intercrosslinked peptides. From the experimental results, both phenyl-based columns are proven to be comparable with the C18 column. Particularly, the fluorophenyl column exhibits significant advantages in identifying important inter-crosslinked peptides and “stacking” the interfering noncrosslinked peptides. These stationary phases are proven suitable to substitute for the commonly used linear alkyl stationary phase in separating and enriching crosslinked peptides, especially their applications in on-line highly sensitive nano-LC-ESI-MS systems.

Conclusion

Although various approaches had been applied for enriching inter-crosslinked peptides for their better identification, it is still necessary to improve and optimize the existing enrichment methods. Two commercialized columns (biphenyl and fluorophenyl), which provided different selectivities in separating polar and hydrophilic compounds, were compared with a versatile C18 column in separating the protein digests of BS³ crosslinked BSA. It was shown that fluorophenyl column has extended retention towards the inter-crosslinked peptides. Through the analysis of fractions from three columns, it is shown that fluorophenyl column increased inter-crosslinked peptide identifications while decreased the non-crosslinked peptide identification, possibly because of the additional electrostatic interactions between crosslinked peptides and fluorophenyl stationary phase. The potential of fluorophenyl column as a competitive substitute in chemical crosslinking research was demonstrated. In future, fluorophenyl based capillary columns will be investigated and evaluated in the crosslinking study with higher biological complexity, this will also circumvent the laborious fractionations and possible uncertainty and errors in this process.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement:

We acknowledge the funding from the UT-Arlington start up fund to Dr. Chowdhury. The study was partially funded from the grant from UA5GM113216-01, NIGMS, NIH. We acknowledge the Shimadzu Institute for Advanced Analytical Chemistry at the University of Texas at Arlington for mass spectrometry support. The authors also acknowledge support from the UT systems proteomics networks for an LTQ Velos Pro mass spectrometer. YZB and KAS would like to thank Restek Corporation for their support of this research.

Abbreviations

2D SCX	two-dimensional strong cation exchange
ABC	ammonium bicarbonate
BS³	bis(sulfosuccinimidyl)suberate
BSA	bovine serum albumin
CID	collision induced dissociation
ESI-IT-TOF	electrospray-ion trap-time-of-flight
FA	formic acid
FDR	false discovery rate
HPLC	high-performance liquid chromatography
LC	liquid chromatography
MS	mass spectrometry
LIT	linear ion trap
PBS	phosphate buffer saline
RP	reverse-phase
SEC	size exclusion chromatography
TIC	total ion current

References

- (1). Sinz A Chemical Cross-Linking and Mass Spectrometry for Mapping Three-Dimensional Structures of Proteins and Protein Complexes. *J. Mass Spectrom* 2003, 38 (12), 1225–1237. 10.1002/JMS.559. [PubMed: 14696200]
- (2). Chowdhury SM; Du XX; Tolic N; Wu S; Moore RJ; Mayer MU; Smith RD; Adkins JN Identification of Cross-Linked Peptides after Click-Based Enrichment Using Sequential Collision-Induced Dissociation and Electron Transfer Dissociation Tandem Mass Spectrometry. *Anal. Chem* 2009, 81 (13), 5524–5532. 10.1021/ac900853k. [PubMed: 19496583]
- (3). Sinz A Chemical Cross-Linking and Mass Spectrometry to Map Three-Dimensional Protein Structures and Protein-Protein Interactions. *Mass Spectrom. Rev* 2006, 25 (4), 663–682. 10.1002/mas.20082. [PubMed: 16477643]

- (4). Gingras AC; Gstaiger M; Raught B; Aebersold R Analysis of Protein Complexes Using Mass Spectrometry. *Nat. Rev. Mol. Cell Biol* 2007, 8 (8), 645–654. 10.1038/nrm2208. [PubMed: 17593931]
- (5). Chakrabarty JK; Naik AG; Fessler MB; Munske GR; Chowdhury SM Differential Tandem Mass Spectrometry-Based Cross-Linker: A New Approach for High Confidence in Identifying Protein Cross-Linking. *Anal. Chem* 2016, 88 (20), 10215–10222. 10.1021/acs.analchem.6b02886. [PubMed: 27649375]
- (6). Leitner A Cross-Linking and Other Structural Proteomics Techniques: How Chemistry Is Enabling Mass Spectrometry Applications in Structural Biology. *Chem. Sci* 2016, 7 (8), 4792–4803. 10.1039/c5sc04196a. [PubMed: 30155128]
- (7). Singh P; Panchaud A; Goodlett DR Chemical Cross-Linking and Mass Spectrometry As a Low-Resolution Protein Structure Determination Technique. *Anal. Chem* 2010, 82 (7), 2636–2642. 10.1021/ac1000724. [PubMed: 20210330]
- (8). Tinnefeld V; Venne AS; Sickmann A; Zahedi RP Enrichment of Cross-Linked Peptides Using Charge-Based Fractional Diagonal Chromatography (ChaFRADIC). *J. Proteome Res* 2017, 16 (2), 459–469. 10.1021/acs.jproteome.6b00587. [PubMed: 28054782]
- (9). Young MM; Tang N; Hempel JC; Oshiro CM; Taylor EW; Kuntz ID; Gibson BW; 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 (11), 5802–5806. [PubMed: 10811876]
- (10). Lasker K; Forster F; Bohn S; Walzthoeni T; Villa E; Unverdorben P; Beck F; Aebersold R; Sali A; Baumeister W Molecular Architecture of the 26S Proteasome Holocomplex Determined by an Integrative Approach. *Proc. Natl. Acad. Sci. U. S. A* 2012, 109 (5), 1380–1387. 10.1073/pnas.1120559109. [PubMed: 22307589]
- (11). Herzog F; Kahraman A; Boehringer D; Mak R; Bracher A; Walzthoeni T; Leitner A; Beck M; Hartl FU; Ban N; et al. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science* (80-.) 2012, 337 (6100), 1348–1352. 10.1126/science.1221483.
- (12). Kao A; Randall A; Yang YY; Patel VR; Kandur W; Guan SH; Rychnovsky SD; Baldi P; Huang L Mapping the Structural Topology of the Yeast 19S Proteasomal Regulatory Particle Using Chemical Cross-Linking and Probabilistic Modeling. *Mol. Cell. Proteomics* 2012, 11 (12), 1566–1577. 10.1074/mcp.M112.018374. [PubMed: 22550050]
- (13). Liu F; Rijkers DTS; Post H; Heck AJR Proteome-Wide Profiling of Protein Assemblies by Cross-Linking Mass Spectrometry (Vol 12, Pg 1179, 2015). *Nat. Methods* 2017, 14 (5), 539
- (14). Navare AT; Chavez JD; Zheng CX; Weisbrod CR; Eng JK; Siehnell R; Singh PK; Manoil C; Bruce JE Probing the Protein Interaction Network of *Pseudomonas Aeruginosa* Cells by Chemical Cross-Linking Mass Spectrometry. *Structure* 2015, 23 (4), 762–773. 10.1016/j.str.2015.01.022. [PubMed: 25800553]
- (15). Shi Y; Pellarin R; Fridy PC; Fernandez-Martinez J; Thompson MK; Li YY; Wang QJ; Sali A; Rout MP; Chait BT A Strategy for Dissecting the Architectures of Native Macromolecular Assemblies. *Nat. Methods* 2015, 12 (12), 1135–+. 10.1038/nmeth.3617. [PubMed: 26436480]
- (16). Aufderheide A; Beck F; Stengel F; Hartwig M; Schweitzer A; Pfeifer G; Goldberg AL; Sakata E; Baumeister W; Forster F Structural Characterization of the Interaction of Ubp6 with the 26S Proteasome. *Proc. Natl. Acad. Sci. U. S. A* 2015, 112 (28), 8626–8631. 10.1073/pnas.1510449112. [PubMed: 26130806]
- (17). Petrotchenko EV; Borchers CH Crosslinking Combined with Mass Spectrometry for Structural Proteomics. *Mass Spectrom. Rev* 2010, 29 (6), 862–876. 10.1002/mas.20293. [PubMed: 20730915]
- (18). Schilling B; Row RH; Gibson BW; Guo X; Young MM MS2Assign, Automated Assignment and Nomenclature of Tandem Mass Spectra of Chemically Crosslinked Peptides. *J. Am. Soc. Mass Spectrom* 2003, 14 (8), 834–850. 10.1016/S1044-0305(03)00327-1. [PubMed: 12892908]
- (19). Lu L; Millikin RJ; Solntsev SK; Rolfs Z; Scalf M; Shortreed MR; Smith LM Identification of MS-Cleavable and Noncleavable Chemically Cross-Linked Peptides with MetaMorpheus. *J. Proteome Res* 2018, 17 (7), 2370–2376. 10.1021/acs.jproteome.8b00141. [PubMed: 29793340]

- (20). Du XX; Chowdhury SM; Manes NP; Wu S; Mayer MU; Adkins JN; Anderson GA; Smith RD Xlink-Identifier: An Automated Data Analysis Platform for Confident Identifications of Chemically Cross-Linked Peptides Using Tandem Mass Spectrometry. *J. Proteome Res* 2011, 10 (3), 923–931. 10.1021/pr100848a. [PubMed: 21175198]
- (21). Lima DB; de Lima TB; Balbuena TS; Neves-Ferreira AGC; Barbosa VC; Gozzo FC; Carvalho PC SIM-XL: A Powerful and User-Friendly Tool for Peptide Cross-Linking Analysis. *J. Proteomics* 2015, 129, 51–55. 10.1016/j.jprot.2015.01.013. [PubMed: 25638023]
- (22). Hoopmann MR; Zelter A; Johnson RS; Riffle M; MacCoss MJ; Davis TN; Moritz RL Kojak: Efficient Analysis of Chemically Cross-Linked Protein Complexes. *J. Proteome Res* 2015, 14 (5), 2190–2198. 10.1021/pr501321h. [PubMed: 25812159]
- (23). Gotze M; Pettelkau J; Schaks S; Bosse K; Ihling CH; Krauth F; Fritzsche R; Kuhn U; Sinz A StavroX-A Software for Analyzing Crosslinked Products in Protein Interaction Studies. *J. Am. Soc. Mass Spectrom* 2012, 23 (1), 76–87. 10.1007/s13361-011-0261-2. [PubMed: 22038510]
- (24). Yang B; Wu YJ; Zhu M; Fan SB; Lin JZ; Zhang K; Li S; Chi H; Li YX; Chen HF; et al. Identification of Cross-Linked Peptides from Complex Samples. *Nat. Methods* 2012, 9 (9), 904–+. 10.1038/nmeth.2099. [PubMed: 22772728]
- (25). Zheng CX; Yang L; Hoopmann MR; Eng JK; Tang XT; Weisbrod CR; Bruce JE Cross-Linking Measurements of In Vivo Protein Complex Topologies. *Mol. Cell. Proteomics* 2011, 10 (10).
- (26). Tan D; Li Q; Zhang MJ; Liu C; Ma CY; Zhang P; Ding YH; Fan SB; Tao L; Yang B; et al. Trifunctional Cross-Linker for Mapping Protein-Protein Interaction Networks and Comparing Protein Conformational States. *Elife* 2016, 5
- (27). Leitner A; Reischl R; Walzthoeni T; Herzog F; Bohn S; Forster F; Aebersold R Expanding the Chemical Cross-Linking Toolbox by the Use of Multiple Proteases and Enrichment by Size Exclusion Chromatography. *Mol. Cell. Proteomics* 2012, 11 (3), M111 014126. 10.1074/mcp.M111.014126.
- (28). Fritzsche R; Ihling CH; Gotze M; Sinz A Optimizing the Enrichment of CrossLinked Products for Mass Spectrometric Protein Analysis. *Rapid Commun. Mass Spectrom* 2012, 26 (6), 653–658. 10.1002/rcm.6150. [PubMed: 22328219]
- (29). Buncherd H; Roseboom W; Ghavim B; Du W; de Koning LJ; de Koster CG; de Jong L Isolation of Cross-Linked Peptides by Diagonal Strong Cation Exchange Chromatography for Protein Complex Topology Studies by Peptide Fragment Fingerprinting from Large Sequence Databases. *J Chromatogr A* 2014, 1348, 34–46. 10.1016/j.chroma.2014.04.083. [PubMed: 24819016]
- (30). Appulage DK; Wang EH; Carroll F; Schug KA Automated Screening of Reversed-Phase Stationary Phases for Small-Molecule Separations Using Liquid Chromatography with Mass Spectrometry. *J. Sep. Sci* 2016, 39 (9), 1638–1647. 10.1002/jssc.201600131. [PubMed: 26959840]
- (31). Buszewski B; Noga S Hydrophilic Interaction Liquid Chromatography (HILIC)-a Powerful Separation Technique. *Anal. Bioanal. Chem* 2012, 402 (1), 231–247. 10.1007/s00216-011-5308-5. [PubMed: 21879300]
- (32). Jardim ICSF; Maldaner L; Lourenco J; Fioravanti LMA; Collins CH Some New Selective Stationary Phases for RP-HPLC. *J. Sep. Sci* 2010, 33 (19), 2917–2929. 10.1002/jssc.201000313. [PubMed: 21038456]
- (33). Brindle R; Klaus A Stationary Phases with Chemically Bonded Fluorene Ligands: A New Approach for Environmental Analysis of Pi-Electron Containing Solutes. *J. Chromatogr. A* 1997, 757 (1–2), 3–20.
- (34). Shollenberger D; Cramer H; Bell DS Evaluation of Retention and Selectivity Using Biphenyl Stationary Phases. *Lc Gc Eur* 2017, 30 (6), 314–320.
- (35). Grassetti AV; Hards R; Gerber SA Offline Pentafluorophenyl (PFP)-RP Prefractionation as an Alternative to High-PH RP for Comprehensive LC-MS/MS Proteomics and Phosphoproteomics. *Anal. Bioanal. Chem* 2017, 409 (19), 4615–4625. 10.1007/s00216-017-0407-6. [PubMed: 28555341]
- (36). Gilar M; Olivova P; Daly AE; Gebler JC Orthogonality of Separation in TwoDimensional Liquid Chromatography. *Anal. Chem* 2005, 77 (19), 6426–6434. 10.1021/ac050923i. [PubMed: 16194109]

- (37). Yong YS; Chong ETJ; Chen HC; Lee PC; Ling YS A Comparative Study of Pentafluorophenyl and Octadecylsilane Columns in High-Throughput Profiling of Biological Fluids. *J. Chinese Chem. Soc* 2017, 64 (6), 699–710. 10.1002/jccs.201600873.
- (38). Huang BX; Kim HY; Dass C Probing Three-Dimensional Structure of Bovine Serum Albumin by Chemical Cross-Linking and Mass Spectrometry. *J. Am. Soc. Mass Spectrom* 2004, 15 (8), 1237–1247. 10.1016/j.jasms.2004.05.004. [PubMed: 15276171]
- (39). Schmidt R; Sinz A Improved Single-Step Enrichment Methods of Cross-Linked Products for Protein Structure Analysis and Protein Interaction Mapping. *Anal. Bioanal. Chem* 2017, 409 (9), 2393–2400. 10.1007/s00216-017-0185-1. [PubMed: 28083664]
- (40). Oliveros, J. C. Venny. An interactive tool for comparing lists with Venn Diagrams. .

Figure 1. Chromatograms of crosslinked BSA sample in (a) C18 column, (b) biphenyl column and (c) fluorophenyl column for 3 replicates (Pink: replicate1, Orange: replicate2, Black: replicate3).

Figure 2.

The matching of b and y ions of inter-crosslinked peptide CCTKPESER(4)-SLGKVGTR(4) from regular LC-ESI-MS/MS (top spectrum) and nano LC-ESI-MS/MS (bottom spectrum), extracted from pFind software.

Figure 3.

Total number of inter-crosslinked peptides identified by three columns after collecting fractions with different parameters and filtering (a. search with 100 ppm precursor and fragment ion tolerance; b. search with 1000 ppm precursor and fragment ion tolerance; c. search with 1000 ppm precursor and fragment ion tolerance with re-filtering with confident parameters in pLink).

Table 1.

Number of inter-crosslinked peptides identified under different concentrations

Column	Number of inter-crosslinked identification		
	0.6 µg/µL	1.7 µg/µL	5.0 µg/µL
C18	1.6±0.7	5.4±1.1	3.8±0.8
Biphenyl	1.8±0.4	5.1±0.9	3.1±0.8
Fluorophenyl	1.6±0.5	4.4±1.4	2.7±0.5
Total IDs	2	10	6

Table 2.

Retention time comparison of the identified inter-crosslinked peptides (N/A implies that the peptide was not detected in certain column)

Inter-crosslinked peptides		Biphenyl	C18	Fluorophenyl
CC-SL	CCTKPESER(4)-SLGKVGTR(4):0	13.3 min	13.5 min	16.0 min
CC-LS	CCTKPESER(4)-LSQKFPK(4):0	14.6 min	14.6 min	17.3 min
LA-AL	LAKEYEATLEECCA(3)-ALKAWSVAR(3):0	20.6 min	20.8 min	24.9 min
LC-CA	LCVLHEKTPVSEK(7)-CASIQKFGGER(6):0	18.0 min	18.6 min	21.9 min
LF-KQ	LFTFHADICTLPDTEKQIK(16)-KQTALVELLK(1):0	24.7 min	25.8 min	N/A
RD-FK	RDTHKSEIAHR(5)-FKDLGEEHFK(2):0	18.0 min	18.6 min	22.0 min
VH-AL	VHKECCHGDLLECADDR(3)-ALKAWSVAR(3):0	17.8 min	N/A	N/A
FK-DT	FKDLGEEHFK(2)-DTHKSEIAHR(4):0	15.0 min	15.8 min	18.9 min
KV-HK	KVPQVSTPTLVEVSR(1)-HKPKATEEQLK(4):0	N/A	N/A	20.0 min
DT-LS	DTHKSEIAHR(4)-LSQKFPK(4):0	N/A	14.9 min	N/A

Table 3.

Combined results of inter-crosslinked peptide and their crosslinked distances. Plausible crosslinked distances are bolded.

Column	Inter-crosslinked peptides	Crosslinked distance(Å)
BP	CCTKPESER(4)-SLGKVGTR(4):0	13.4
BP	LAKEYEATLEECCAK(3)-ALKAWSVAR(3):0	13.6
BP	GACLLPKIETMR(7)-SLGKVGTR(4):0	19.7
BP	CCTKPESER(4)-LSQKFPK(4):0	20.2
BP	CCTKPESER(4)-PKATEEQLK(2):0	35.3
BP	PKATEEQLK(2)-LSQKFPK(4):0	47.1
BP	KFWGK(1)-HKPK(2):0	51.7
BP	PKATEEQLK(2)-KFWGK(1):0	54.4
BP	PKATEEQLK(2)-LKECCDK(2):0	58.5
C18	CCTKPESER(4)-SLGKVGTR(4):0	13.4
C18	LAKEYEATLEECCAK(3)-ALKAWSVAR(3):0	13.6
C18	CCTKPESER(4)-LSQKFPK(4):0	20.2
C18	TPVSEKVT(6)-LSQKFPK(4):0	35.5
C18	AEFVEVTKLVTDLT(8)-HKPK(2):0	55.1
C18	PKATEEQLK(2)-LKECCDK(2):0	58.5
FP	CCTKPESER(4)-SLGKVGTR(4):0	13.4
FP	LAKEYEATLEECCAK(3)-ALKAWSVAR(3):0	13.6
FP	NYQEAKDAFLGSFLYEYSR(6)-ALKAWSVAR(3):0	14.4
FP	ALKAWSVARLSQK(3)-LSQKFPK(4):0	15.1
FP	ALKAWSVAR(3)-LSQKFPK(4):0	15.1
FP	CCTKPESER(4)-LSQKFPK(4):0	20.2
FP	ADEKKFWGK(5)-KFWGK(1):0	27.2
FP	DTHKSEIAHR(4)-ADEKK(4):0	31.3
FP	PKATEEQLK(2)-EKVLASSAR(2):0	38.4
FP	PKATEEQLK(2)-LSQKFPK(4):0	47.1
FP	RDTHKSEIAHR(5)-QIKK(3):0	56.4