HHS Public Access

Author manuscript

J Proteome Res. Author manuscript; available in PMC 2020 September 04.

Published in final edited form as:

J Proteome Res. 2019 June 07; 18(6): 2545–2558. doi:10.1021/acs.jproteome.9b00085.

Optimized Cross-Linking Mass Spectrometry for in Situ Interaction Proteomics

Zheng Ser[†,‡], **Paolo Cifani**[†], **Alex Kentsis**[*,†,‡,§,‖]

[†]Molecular Pharmacology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, New York 10065, United States

[‡]Tri-Institutional Ph.D. Program in Chemical Biology, Memorial Sloan Kettering Cancer Center, New York, New York 10065, United States

[§]Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, New York 10065, United States

[‖]Department of Pediatrics, Pharmacology, and Physiology & Biophysics, Weill Cornell Medical College, Cornell University, New York, New York 10065, United States

Abstract

Recent development of mass spectrometer cleavable protein cross-linkers and algorithms for their spectral identification now permits large-scale cross-linking mass spectrometry (XL-MS). Here, we optimized the use of cleavable disuccinimidyl sulfoxide (DSSO) cross-linker for labeling native protein complexes in live human cells. We applied a generalized linear mixture model to calibrate cross-link peptide-spectra matching (CSM) scores to control the sensitivity and specificity of large-scale XL-MS. Using specific CSM score thresholds to control the false discovery rate, we found that higher-energy collisional dissociation (HCD) and electron transfer dissociation (ETD) can both be effective for large-scale XL-MS protein interaction mapping. We found that the coverage of protein–protein interaction maps is significantly improved through the use of multiple proteases. In addition, the use of focused sample-specific search databases can be used to improve the specificity of cross-linked peptide spectral matching. Application of this approach to human chromatin labeled in live cells recapitulated known and revealed new protein interactions of nucleosomes and other chromatin-associated complexes in situ. This optimized approach for mapping native protein interactions should be useful for a wide range of biological problems.

[*]**Corresponding Author** kentsisresearchgroup@gmail.com.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.9b00085.

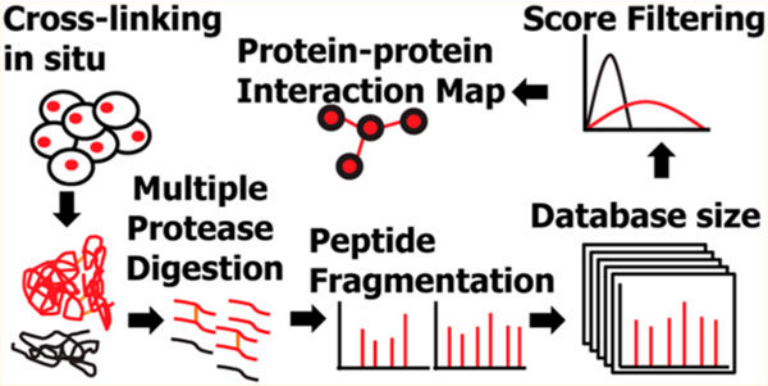
Figure S1: Cross-linked peptide properties across fragmentation methods; Figure S2: Cross-linked peptide properties across different proteases; Figure S3: Comparison of cross-links and cross-link score with published data (PDF)

Table S1: List of cross-linked chromatin proteins (XLSX)

The authors declare the following competing financial interest(s): A.K. is a consultant for Novartis. Other authors declare no competing financial interest.

Mass spectrometry raw files and search results are publicly available through the ProteomeXchange data repository via the PRIDE database with the data set identifier PXD010796.

Graphical Abstract



Keywords

cross-linking; mass spectrometry; chromatin; BSA; protein–protein interactions; target-decoy strategy; database search; false positive discovery; proteomics

INTRODUCTION

Protein–protein interactions mediate cellular functions including signaling, metabolism, and cell differentiation. Disruption of physiologic protein–protein interactions can contribute to disease, and defining how protein complexes change can provide important insights into pathogenic mechanisms, and ultimately lead to improved therapies.[1] Conventional methods to study protein–protein interactions are based on large-scale affinity purification or separation of protein complexes combined with protein identification using mass spectrometry.[2–6] These methods are typically performed under nonphysiological conditions, such as ectopic protein expression or use of affinity tags to facilitate purification of the complexes, which can produce artifacts and disrupt native protein interactions. In contrast, cross-linking mass spectrometry (XL-MS) relies on covalent labeling and cross-linking of proteins in physical proximity, thereby identifying native protein–protein interactions. This approach can also identify transient interactions that occur dynamically during specific cell or developmental states.[7,8] In addition, high-resolution mass spectrometry can directly identify sites of interaction between cross-linked amino acids, providing detailed spatial information for structural studies.[9,10] Recently, development of collision-induced dissociation (CID) cleavable cross-linkers[11–15] and improvements in algorithms for cross-link peptide spectral analysis[16–20] has enabled protein–protein interaction mapping of complex proteomes, including worm,[21,22] bacteria,[22–24] and human,[16,17,25,26] and various other organisms,[27] with hundreds-to-thousands of cross-links and protein–protein interactions identified.[17,22] While conventional cross-linking is typically performed on cell extracts or purified proteins,[16,17,21,22] the membrane permeability of some mass spectrometer cleavable cross-linking reagents enables cross-linking of protein complexes in live cells to define specific protein–protein interactions in situ.[14,28]

To facilitate large-scale interaction proteomics, we sought to define the optimal methods for the isolation of cross-linked peptides from protein complexes labeled in situ, and to ensure their accurate identification using a generalized linear mixture model to discriminate between true and false identifications for cross-linking mass spectrometry. Thus, we present optimized parameters for the use of the CID-cleavable cross-linker disuccinimidyl sulfoxide (DSSO) in situ. We found that higher-energy collisional dissociation (HCD) fragmentation produced as many cross-linked protein complex identifications as electron transfer dissociation (ETD) fragmentation using either whole proteome or focused target databases for cross-link peptide spectral matching. The use of multiple proteases significantly increased the coverage of protein–protein interactions maps. Importantly, the developed generalized linear mixture model strategy allowed for explicit control of false discovery and optimization of sensitivity and specificity of protein interactions maps in cells in situ. Application of this approach to human chromatin labeled in live cells recapitulated known and revealed new protein interactions of nucleosomes and other chromatin-associated complexes.

EXPERIMENTAL SECTION

Reagents

Disuccinimidyl sulfoxide (DSSO), formic acid, dimethyl sulfoxide (DMSO), and LC–MS grade solvents were obtained from Thermo Scientific. Bovine serum albumin (BSA), (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), sodium chloride, magnesium chloride, potassium chloride, sucrose, ethylenediaminetetraacetic acid (EDTA), dithiothreitol, guanidinium hydrochloride, *Staphylococcus aureus* micrococcal nuclease, iodoacetamide, ammonium bicarbonate and formic acid were obtained from Millipore Sigma. Protease inhibitors 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF) and pepstatin were obtained from Santa Cruz, bestatin from Alfa Aesar, and leupeptin from EMD Millipore. Trypsin and GluC proteases were obtained from Promega. Chymotrypsin protease was obtained from Pierce Thermo. LysC was obtained from Wako Pure Chemical Industries.

Cell Culture and Protein Isolation

HEK293T cells were cultured in DMEM supplemented with 10% (v/v) fetal bovine serum, penicillin and streptomycin. HEK293T cells (50 million) were sedimented by centrifugation at 500*g* and lysed in 500 μL 6 M guanidine hydrochloride in 100 mM ammonium bicarbonate buffer, pH 8. The lysate was sonicated (E210 adaptive focused acoustic sonicator, Covaris) for 5 min at 4 °C and homogenized by passing through a 27G needle. The lysate was then clarified by centrifugation at 18 000*g* for 10 min at 4 °C. Protein concentration of clarified lysates was determined using the bicinchoninic acid (BCA) assay, according to manufacturer's instructions (Thermo Scientific).

Preparation of Cross-Linked Bovine Serum Albumin and Peptide Purification

50 μg of BSA was solubilized in 50 μL HEPES buffer (20 mM HEPES, pH 8, 150 mM sodium chloride, 1.5 mM magnesium chloride, 0.5 mM dithiothreitol) and cross-linked by adding 1 μL of 50 mM DSSO dissolved in DMSO (1:1 protein amine:DSSO molar ratio).

The solution was incubated at room temperature for 1 h, before quenching by the addition of 1 M ammonium bicarbonate buffer to a final concentration of 20 mM. The cross-linked protein was denatured by adding 50 μL of 6 M guanidine hydrochloride in 100 mM ammonium bicarbonate, pH 8, then reduced by adding 11 μL of 100 mM dithiothreitol at 56 °C for 1 h and alkylated with 12 μL of 550 mM iodoacetamide at room temperature for 30 min in the dark. Alkylation was quenched by the addition of 13 μL of 100 mM dithiothreitol and incubated at 56 °C for 20 min. The solution was diluted with 50 mM ammonium bicarbonate solution, pH 8, for a final concentration of 0.2 M guanidine hydrochloride. For digestion with a single protease, either trypsin, chymotrypsin or GluC was added in 1:50 (enzyme:protein, w/w) concentration and incubated at 37 °C for 18 h. For sequential digestion with LysC followed by trypsin, LysC was first added in 1:100 (enzyme:protein, w/w) concentration and incubated at 37 °C for 8 h. Trypsin was then added in 1:50 (enzyme:protein, w/w) concentration and incubated at 37 °C for 18 h. Digestion was stopped by the addition of formic acid to 1% (v/v) concentration. Peptides were then purified using solid phase extraction with C18 Macrospin columns according to manufacturer's instructions (Nest Group), and concentrated by vacuum centrifugation.

Protein Cross-Linking of Live Cells in Situ and Isolation of Cross-Linked Peptides from Chromatin

HEK293T cells (50 million) were sedimented by centrifugation at 500*g* and then suspended in 1 mL of hypotonic buffer (10 mM HEPES, pH 8, 10 mM sodium chloride, 1 mM magnesium chloride, 0.5 mM dithiothreitol, protease inhibitors). For cross-linking, 50 μL of 50 mM DSSO dissolved in DMSO was added and cells were incubated for 1 h at 4 °C. The reaction was then quenched by the addition of 50 μL of 500 mM Tris-HCl, pH 8. Cells were subsequently subjected to 15 strokes of Dounce homogenization and the solution was then centrifuged for 15 min at 3300*g* to isolate nuclei. Sedimented nuclei were then resuspended in 500 μL of 0.25 M sucrose, 10 mM magnesium chloride buffer and layered on top of a solution of 500 μL of 0.88 M sucrose, 0.05 mM magnesium chloride. The resultant mixture was centrifuged for 10 min at 1200*g*. Purified nuclei were then resuspended in 1 mL of 10 mM HEPES, pH 7.4, 0.2 mM magnesium chloride solution and incubated on ice for 10 min to extract nucleoplasm. The solution was then centrifuged for 20 min at 770*g*. The pellet was next resuspended in 1 mL of 10 mM HEPES, pH 7.4, 500 mM sodium chloride, 0.2 mM magnesium chloride and incubated on ice for 10 min to extract bound proteins. The solution was then centrifuged for 20 min at 770*g*. Next, the pellet was solubilized in 500 μL of detergent buffer (15 mM HEPES, pH 7.5, 500 mM sodium chloride, 1 mM dithiothreitol, 0.34 M sucrose, 10% glycerol, 0.5% v/v Triton X-100) and incubated on ice for 10 min. The solution was then layered on top of a 500 μL solution of 0.88 M sucrose, 0.05 mM magnesium chloride and centrifuged for 20 min at 3300*g* to purify chromatin. The chromatin pellet was resuspended in 500 μL of 15 mM HEPES, pH 7.5, 15 mM sodium chloride, 60 mM potassium chloride, 5 mM magnesium chloride, 1 mM calcium chloride, 0.25 M sucrose and protease inhibitors (0.5 M AEBSF, 0.001 mM pepstatin, 0.01 mM bestatin and 0.1 mM leupeptin). The chromatin suspension was incubated at 37 °C for 10 min. Five units of micrococcal nuclease were subsequently added and incubated at 37 °C for 1 h. The nuclease digestion was stopped by the addition of EDTA to a final concentration of

50 mM and the solution was centrifuged for 10 min at 18 000*g* to remove the nuclear matrix. The supernatant containing cross-linked chromatin was then quantified using the BCA assay.

Cross-linked chromatin proteins were then purified by filter-aided sample preparation (FASP) using Sartorius Vivacon 500 with a 30 000 MW cutoff filter, as previously described.[29,30] Briefly, 200 μg of protein was diluted 5-fold in urea buffer (8 M urea, 50 mM ammonium bicarbonate, pH 8.0) and applied to the filter unit. Filter units were centrifuged at 14 000*g* for 20 min. The membrane was then washed twice with 200 μL urea buffer by centrifugation at 14 000*g* for 20 min. The membrane was then incubated with 200 μL of 100 mM dithiothreitol in urea buffer for 20 min at room temperature before centrifugation at 14 000*g* for 20 min. The proteins were then incubated with 200 μL of 100 mM iodoacetamide in urea buffer for 20 min at room temperature in the dark before centrifugation at 14 000*g* for 20 min. The membrane was washed twice more with urea buffer then once with 200 μL of 50 mM ammonium bicarbonate, pH 8. Proteins retained on the membrane were then digested with trypsin at 1:50 (enzyme:protein, w/w) concentration at 37 °C for 18 h, or with lysC at 1:100 (enzyme:protein, w/w) at 37 °C for 8 h followed by trypsin digestion or with either chymotrypsin or gluC at 1:20 (enzyme:protein, w/w) concentration at 37 °C for 18 h. Digested peptides were collected by centrifugation at 14 000*g* for 20 min. Peptides were then purified using solid phase extraction with C18 Macrospin columns according to manufacturer's instructions (Nest Group). The purified peptides were concentrated by vacuum centrifugation and resuspended in 100 μL of 5% acetonitrile, 0.1% formic acid. The peptides were then injected onto the strong cation exchange (SCX) column (Protein Pak Hi-Res SP 7 μm 4.6 × 100 mm, Water) at a flow rate of 0.5 mL/min with a column temperature of 30 °C using the Alliance HPLC system (Waters). Gradient was run with 5% acetonitrile and 0.1% formic acid in water for solvent A and 1 M potassium chloride in 5% acetonitrile and 0.1% formic acid in water for solvent B. Column gradient was set as follows: 0–3 min (0% B); 3–33 min (0–10% B); 33–43 min (10–100% B); 43–60 min (100% B); 60–70 min (100−0% B); 70–90 min (0% B). Six fractions were collected from 38 to 60 min for trypsin and lysC followed by trypsin digestions, while five fractions were collected from 38 to 50 min for chymotrypsin and gluC digestions. Fractions were then purified using solid phase extraction with C18 Macrospin columns according to manufacturer's instructions (Nest Group). Peptides were concentrated by vacuum centrifugation and then stored at −20 °C before analysis.

Liquid Chromatography and Nanoelectrospray Mass Spectrometry

Cross-linked peptides were separated by reverse phase nanoflow liquid chromatography (EKspert nanoLC 425, Ekisgent) coupled to the Orbitrap Fusion mass spectrometer (Thermo). Cross-linked BSA peptides and HEK293T peptides were resuspended in 0.1% formic acid in water (v/v). For cross-linked BSA peptides, 1 μg of peptide was injected for analysis. For cross-linked BSA peptides mixed with non-cross-linked HEK293T peptides in 1:1 (w/w) ratio, 1 μg of total peptide was analyzed. For SCX fractions of cross-linked chromatin, peptides were resuspended in 20 μL of 0.1% formic acid in water (v/v) and 0.5 μg of peptides were analyzed.

For liquid chromatography, a trap-elute system was used, as described previously.[31] Reverse phase columns were fabricated as previously described.[32] Trap columns were fabricated by packing Poros R2 10 μm C18 particles (Life Technologies) into 4 cm fritted capillaries with internal diameters of 150 μm. Reverse phase columns were fabricated by packing Reprosil 1.9 μm silica C18 particles into 40 cm fritted capillaries with internal diameters of 75 μm. Peptides were resolved over a 90 min gradient from 5% to 40% acetonitrile and ionized using the DPV-565 Picoview ion source (New Objective) operated with the ionization voltage of 1700 V. Custom electrospray emitters were fabricated as previously described.[33]

For mass spectrometry, five different acquisition methods were used: (1) CID-MS2/HCD-MS2, (2) CID-MS2/ETD-MS2, (3) CID-MS2/EThcD-MS2, (4) CID-MS2/HCD-MS3, (5) CID-MS2-MS3/ETD-MS2 where HCD is the higher-energy collisional dissociation, ETD is electron transfer dissociation, and EThcD is the hybrid electron transfer with supplemental activation by higher energy collision dissociation. For MS2-MS2 methods, precursor ion spectra were recorded at 400–1800 m/z with 60 000 m/z Orbitrap resolution, automatic gain control target of 1×10^5 ions and maximum injection time of 50 ms. Precursor ions with 3–10 positive charge were selected for MS2 fragmentation with dynamic exclusion of 60 s after 1 scan and isolation window of 2 Th. For CID-MS2, precursor ion spectra were recorded in the Orbitrap with resolution 30 000 m/z, automatic gain control target of 5.0×10^4 and maximum injection time of 100 ms at normalized collision energy of 30%. For HCD-MS2, precursor ion spectra were recorded in the Orbitrap with resolution 30 000 m/z, automatic gain control target of 5.0×10^4 and maximum injection time of 120 ms at normalized collision energy of 30%. For ETD-MS2, precursor ion spectra were recorded in the Orbitrap with resolution 30 000 m/z, automatic gain control target of 5.0 × 104 ions and maximum injection time of 120 ms with fluoroanthene reaction time of 50 ms, ETD target of 2.0×10^5 ions and maximum fluoroanthene injection time of 200 ms. For EThcD-MS2, precursor ion spectra were recorded in the Orbitrap with resolution 30 000 m/z, automatic gain control target of 5.0×10^4 ions and maximum injection time of 120 ms with reaction time of 50 ms, reagent target of 2.0×10^4 and maximum fluoroanthene injection time of 200 ms and supplemental collision energy of 20%. For HCD-MS3, MS2 fragment ions with 2–6 charge states and mass difference of 31.9721 Da with 1–100% intensity range were selected for MS3 fragmentation, with a mass tolerance of 30 ppm and 2 Th isolation window. For CID-MS2-MS3, MS2 fragment ions with mass difference of 31.9721 Da with 1–100% intensity range were selected for MS3 fragmentation, with a 3 Th isolation window and normalized collision energy of 35%. MS3 fragment ion spectra were recorded in the linear ion trap acquisition with automatic gain control target of 2.0×10^4 ions and maximum injection time of 150 ms.

Data Analysis

Acquired mass spectra raw files were analyzed using XlinkX version 2.2,[16,17] as implemented in Proteome Discoverer version 2.2. For CID-MS2-MS3/ETD-MS2 fragmentation, separate searches were performed for MS2 and MS3 identification of cross-linked peptides and the results from both MS2 and MS3 identification were merged for subsequent analysis. The target search database was the *Homo sapiens* proteome from SwissProt, containing isoforms[34] (version January 2016), supplemented with the sequences

of common contaminant proteins from cRAP.[35] Mass spectral analysis parameters included 2 maximum allowed missed cleavages, minimum peptide length of 5 and 4 maximum variable modifications: methionine oxidation, hydrolysis of lysyl-DSSO by water, and lysyl-DSSO-Tris adduct. Precursor mass tolerance was set at 10 ppm with fragment mass tolerance of 20 ppm and 0.5 Da for Orbitrap and ion trap spectra, respectively. Cysteine carbamidomethylation was set as fixed modification. A 1% false discovery rate (FDR) was set for cross-link peptide spectra matching with minimum score threshold set at 0 using the Percolator algorithm.[36] Peptide spectral matches and their scores were analyzed using R scripts to concatenate cross-links identified from biological replicates (https://github.com/kentsisresearchgroup/R-script-for-XLMS-data-analysis). For histone proteins, peptides that could be identified as exclusively belonging to specific isoforms were marked with their specific isoform protein name or gene name. Proteins that could be mapped to multiple isoforms were labeled with the family protein or gene name, e.g., peptides that could belong to histone isoforms H3.1, H3.2 or H3.3 were labeled as histone H3 peptides. Bar graphs and scatter plots were plotted using Origin 2018 (Microcal). Venn diagrams were made with Venny version 2.1.[37] Intersection plots were made using UpSetR App.[38] Peptide cross-link sequences were visualized using xiNET.[39] Atomic resolution structures were visualized with UCSF Chimera[40] version 1.12 and cross-links were mapped using Xlink analyzer.[41] Protein–protein interaction maps were made with Cytoscape version 3.6.0.[42] Mass spectrometry raw files and search results are publicly available through the ProteomeXchange data repository via the PRIDE database with the data set identifier PXD010796.

RESULTS AND DISCUSSION

Calibration of Cross-Link Peptide-Spectra Matching (CSM) Scores To Control the Sensitivity and Specificity of Large-Scale XL-MS

MS-cleavable cross-linkers are designed to improve the accuracy and sensitivity of identification of protein–protein interactions by mass spectrometric analysis and subsequent statistical spectral matching. Specifically, the DSSO cross-linker is labile upon collision-induced dissociation, permitting the uncoupling of cross-linked peptides in gas phase and thus enabling the independent identification of each individual peptide of the cross-linked peptide pair. In addition, fragmentation of the sulfoxide cross-linker leaves characteristic adducts on the labeled residues, thus producing diagnostic mass differences that can be leveraged to improve the specificity of cross-linked peptide identification.[13] Depending on the experimental parameters used, fragmentation of the peptide backbone can be achieved either by fragmenting the cross-linked molecule, where the MS2 spectrum will contain fragments from both cross-linked peptides or fragmenting the decoupled peptide ions, where the decoupled peptides ions can be selected for further MS3 fragmentation. Data generated using both of these MS strategies can be analyzed by recently developed algorithms, which assign scores to matches and utilize the target-decoy approach to discriminate between correctly and incorrectly matched cross-link peptide spectra.[16,17,20,43]

To further advance our ability to control the sensitivity and specificity of cross-linked peptide discovery, we devised a strategy to discriminate between true and false positive

CSMs, based on internal controls to model the score distribution of correct and incorrect cross-link matches. Bovine serum albumin (BSA) is a widely used model protein for mass spectrometry that spontaneously oligomerizes in solution.[44,45] To model a typical XL-MS sample, containing both cross-linked and non-cross-linked peptides, we cross-linked purified BSA, and sequentially digested it with LysC and trypsin proteases. Cross-linked BSA peptides were either analyzed on their own, or diluted with non-cross-linked tryptic peptides from human HEK293T cell lysates (Figure 1A). We reasoned that this mixture model of true positive CSMs (cross-linked BSA peptides) and true negative CSMs (non-cross-linked human peptides) should provide a useful means to control the FDR of large-scale cross-linked mass spectrometry. To test this idea, we implemented a generalized linear mixture model to assess the specificity of cross-linked peptide identification from data recorded using five different fragmentation strategies: CID-MS2/HCD-MS2, CID-MS2/ETD-MS2, CID-MS2/EThcD-MS2, CID-MS2/HCD-MS3, and CID-MS2-MS3/ETD-MS2 (Figure 1B). Using the XlinkX algorithm,[16,17] we scored the MS spectra against a human reference database containing BSA and other contaminant proteins and applied a 1% FDR threshold for peptide-spectral matching using Percolator.[36,46] Here, we observed that a substantial number of incorrect CSMs passed the 1% Percolator FDR for both cross-linked BSA and cross-linked BSA diluted with non-cross-linked human proteome (Figure 1B). For example, for CID-MS2/HCD-MS2 fragmentation, 173 of 291 (59%) of cross-links were mapped to human proteins, indicating incorrect CSMs, even after controlling the mass spectral assignment FDR by Percolator. Manual inspection of representative spectra confirmed that these apparent incorrect matches were due either to incomplete fragmentation of one of the peptides in the cross-link pair or relatively low signal intensity of the diagnostic DSSO fragment ions. We ruled out the possibility of incorrect cross-link matches due to the misassignment of non-cross-linked peptide spectra as cross-linked spectra, since none of the examined spectra from apparently incorrect matches exhibited fragment ions consistent with albumin peptide sequences. In addition, we also verified that none of the mismatched spectra were due to the incomplete DSSO cross-linking reaction due to spectra with single site DSSO adducts. Lastly, we also excluded the possibility that incorrect CSMs were produced by cross-linking of short nonproteotypic peptides. We observed that both cross-linked BSA peptides and non-BSA cross-linked peptides share similar peptide lengths, and the shorter peptides in identified cross-links (marked as peptide A in Figure S1) have an average length of approximately 10 residues.

Next, we sought to improve the accuracy of cross-link identification using the CSM score distributions of the true positive CSMs and true negative CSMs to define empiric thresholds for CSM filtering based on the CSM score as calculated using the XlinkX algorithm (Figure 1B).[16,17] Indeed, as we fit the observed CSM score distributions to a linear mixture of two Gaussian functions, we observed two CSM score modes corresponding to apparently true and false matches (Figure 1C). On the basis of these distributions, we expressed the fraction of incorrect CSMs (non-cross-linked human peptides) as a function of the score. Specifically, we calculated that a CSM score threshold of 100 corresponded to 1% incorrect cross-link identification (Figure 1D), and we defined a CSM score threshold of 74 for improved sensitivity at a more relaxed (10%) error rate. Importantly, the use of empiric CSM score thresholds reduced the number of incorrect cross-links (i.e., cross-links on human

proteins) from 58 to 3 (Figure 1E). To further confirm the improved specificity achieved using the empirically defined threshold, we mapped the identified BSA cross-links to the high-resolution atomic-level structure of BSA (PDB ID 4f5s).[47] Consistently, the use of CSM score thresholds eliminated cross-links between lysine residues that are farther than 26.0 Å, the upper limit of the Cα–Cα distance afforded by the DSSO cross-linker,[13] while retaining cross-links within the expected lysyl-DSSO distance constraints (Figure 1F).

These results demonstrate the need to control the false identifications inherent in the statistical spectral matching for large-scale cross-linking mass spectrometry proteomics beyond target-decoy strategies and provide a facile means to do so using a score threshold strategy adapted for cross-linking mass spectrometry. Ideally, true and false cross-link mixtures should be explicitly incorporated into experimental design. However, in the absence of such direct comparisons, CSM score thresholds established for closely related mixtures, as based on the similarity of peptide composition, length, and ion charge states, can provide approximations (Figure S1).

Improved Cross-Linked Peptide Identification Using CID-MS2/HCD-MS2 Fragmentation

Comparison of CID, HCD, and ETD fragmentation of linear non-cross-linked peptides indicates that the efficiency of each fragmentation method to produce optimal fragment ion spectra depends on the specific physicochemical properties of fragmented peptides.[48] For example, ETD exhibits reduced fragmentation efficiency for larger m/z peptides.[49] Previous studies of noncleavable cross-linkers found that HCD fragmentation increased the identification of cross-linked peptides with charge states of 3–4, while the hybrid EThcD method facilitated the identification of cross-linked peptides with charge states of ≥5.[50] Recent studies of MS-cleavable cross-linkers identified multiple fragmentation modes as integral to increasing the cross-link peptide identification.[17,51]

Thus, to assess the efficiency of various fragmentation strategies for the identification of DSSO cross-linked peptides, we evaluated the sensitivity of cross-link identification using the mixture of cross-linked BSA and non-cross-linked human proteome, sequentially digested by LysC followed by trypsin, and applied the empiric CSM score thresholds. We compared 5 different fragmentation methods, with CID-MS2 used for cross-link cleavage: CID-MS2/HCD-MS2, CID-MS2/ETD-MS2, CID-MS2/EThcD-MS2, CID-MS2/HCD-MS3, and CID-MS2-MS3/ETD-MS2 (Figure S1). We observed an average of 43 cross-links per experiment across 3 biological replicates identified by CID-MS2/HCD-MS2, as compared to 25, 19, 7, and 19 cross-links identified with CID-MS2/ETD-MS2, CID-MS2/HEThcD-MS2, CID-MS2/HCDMS3, and CID-MS2-MS3/ETD-MS2 methods, respectively (Figure 2A). A total of 57 unique cross-links across 3 biological replicates were identified using CID-MS2/HCD-MS2 fragmentation, while 41 unique cross-links were identified using CID-MS2/ETD-MS2. Out of these unique cross-links, only 4 were identified by CID-MS2/ETD-MS2 but not by CID-MS2/HCD-MS2 (Figure 2B). For cross-linked BSA peptides mixed with the non-cross-linked proteome, an average of 21 cross-links per experiment across 3 biological replicates were identified by CID-MS2/HCD-MS2, as compared to 19, 13, 3, and 12 cross-links identified by CID-MS2/ETD-MS2, CID-MS2/HEThcD-MS2, CID-MS2/HCDMS3, and CID-MS2-MS3/ETD-MS2 methods, respectively (Figure 2C). Thirty one unique BSA

cross-links across 3 biological replicates were identified from CID-MS2/HCD-MS2, while 27 were identified from CID-MS2/ETD-MS2. CID-MS2/HCD-MS2, and CID-MS2/ETD-MS2 identified 9 and 3 cross-linked peptides, respectively, that were not identified by other fragmentation methods, suggesting that these two fragmentation methods are complementary to each other (Figure 2D). The CID-MS2/HCD-MS2 and CID-MS2/ETD-MS2 methods showed a large overlap of identified cross-links, with 32 out of 62 (59%) and 18 out of 40 (45%) shared cross-linked BSA peptides identified in pure cross-linked BSA sample or with background human proteome, respectively. Interestingly, the CID-MS2-MS3/ETD-MS2 method identified as many BSA cross-links as the CID-MS2/EThcD-MS2 method and fewer BSA cross-links as compared to fragmentation by CID-MS2/ETD-MS2 and CID-MS2/HCD-MS2. Consistent with prior reports, CID-MS2-MS3/ETD-MS2 identifies at least as many number of total unique cross-links for both BSA and non-BSA cross-links, as CID-MS2/ETD-MS2 (90 cross-links) before the application of empiric CSM score thresholds.[17] This suggests that the specific CSM scoring algorithm for MS2-MS2 spectra compared to MS2-MS3-MS2 results in different stringency when applying the CSM score thresholds.[16,17] Combined with the longer acquisition times required by the CID-MS2-MS3/ETD-MS2 method, this might explain the reduced sensitivity of cross-link identification for complex samples.

Comparison of identified cross-linked peptides showed relatively shorter peptide length, lower precursor ion m/z and lower cross-linked peptide identifications across all charge states obtained by ETD fragmentation, as compared to that obtained by HCD fragmentation (Figure S1), in agreement with prior observations.[49] A small fraction (10–30%) of cross-linked peptides were identified exclusively by either ETD or HCD fragmentation methods (Figure 2B,D), consistent with the complementarity of these fragmentation methods.

In all, these results indicate that CID-MS2/HCD-MS2 fragmentation yields the highest number of cross-links for simple samples, with a 60% increase in unique cross-link sites as compared to CID-MS2/ETD-MS2 (Figure 2A,B). For samples with a proteomic background, CID-MS2/HCD-MS2 identified as many unique cross-link sites as CID-MS2/ETD-MS2 (Figure 2C,D). On the other hand, CID-MS2/HCD-MS3 and CID-MS2/EThcD-MS2 methods consistently recorded fewer cross-link spectra and fewer identifications when compared to CID-MS2/ETD-MS2 and CID-MS2/HCD-MS2 methods (Figure 2A–D). This can be attributed to the relatively longer MS duty-cycle, and consequently lower sampling rate, imposed by MS3 and EThcD methods, as also noted in a recent study.[17] Given the relatively wide range of physical and chemical properties of cross-linked BSA peptides, these results should be generalizable to other proteomes.

Focused Target Databases for Identification of Protein–Protein Interactions by XL-MS

For noncleavable cross-linkers, individual peptides of a cross-link pair have to be searched combinatorially, exponentially increasing the database search space and computing times.[52] In contrast, mass spectrometer cleavable cross-linkers permit database searching of individual peptide of the cross-link pair, reducing the database search space and time, which allows the use of proteome wide databases when compared to noncleavable cross-linkers.[16] We thus hypothesized that reducing the search space using focused databases can further

improve the sensitivity of mass spectral matching,[53–56] but may affect the accuracy of identifications due to the insufficient number of decoy sequences.[57,58] Since cross-linking mass spectrometry is often used for the analysis of specific purified protein complexes, we examined the effects of using a focused database as compared to a complete reference proteome on mass spectrometric identification of DSSO cross-linked peptides.[52] To this end, we generated search databases containing the BSA sequence alone, or BSA plus the human reference proteome supplemented with common contaminants from cRAP.[35] We then used these databases as targets to score recorded spectra by analyzing cross-linked BSA peptides or cross-linked BSA peptides diluted in non-cross-linked human proteome. From spectra recorded without background, we identified an average of 43 BSA cross-links across 3 biological replicates when searching against either the BSA-specific or reference human proteome databases (Figure 2E). For cross-linked BSA peptides mixed with the non-cross-linked proteome, we found an average of 21 and 22 BSA cross-links from 3 biological replicates when searching against the BSA-specific database or BSA plus reference human proteome databases, respectively (Figure 2E). This suggests that the database size and composition do not affect the sensitivity of cross-link identification per se (Figure 2F). However, the use of focused database did not increase the number of BSA cross-links identified as a result of incorrect matching of spectra from non-cross-linked peptide. This indicates that the use of focused databases does not reduce specificity of cross-link identification. In addition, the use of focused databases permits the inclusion of variable chemical modifications that would be required for the detection of post-translational modifications, which would otherwise be precluded by the current computational requirements of proteome-wide analysis.

Diverse Sequence-Specific Proteases To Expand the Coverage and Density of DSSO XL-MS Protein Interaction Maps

We reasoned that the DSSO reaction with lysine residues reduces the accessibility of cross-linked peptides for tryptic proteolysis, thus reducing the identification of tryptic cross-linked peptides. Previous studies reported the use of Arg-N or sequential digestion using trypsin and elastase to improve cross-links identification as compared to conventional tryptic digestion.[59,60] For linear non-cross-linked peptides, additional sequence-specific proteases are known to identify additional peptide sequences, thus improving protein sequence coverage.[61–63] To test this hypothesis for XL-MS, we utilized the model containing cross-linked BSA peptides diluted in non-cross-linked proteins extracted from human HEK293T cells (Figure 1A). Cross-linked BSA protein was digested (1) with trypsin only, (2) sequentially with LysC followed by trypsin, (3) with GluC only, or (4) with chymotrypsin only. The cross-linked BSA peptides were diluted into proteome lysates digested with the same proteases. We observed a substantial number of apparently incorrectly matched cross-linked peptide spectra for GluC and chymotrypsin digests. Similarly to what was observed using tryptic proteolysis, the CSM score distributions of BSA cross-links (true) and non-BSA cross-links (false) were distinct even using GluC and chemotrypsin (Figure 3A). Applying the empiric CSM score threshold approach resulted in cutoffs of 70 and 75 to achieve a 1% cross-link identification error rate for GluC and chymotrypsin digestion, respectively, as compared to 100 obtained for tryptic proteolysis.

Consistent with the previous studies reporting that sequential digestion improves cross-link site identification,[60] we found substantially more unique cross-links identified from the sequential protease digest of cross-linked BSA using LysC endoprotease followed by trypsin, as compared to that of trypsin alone (22 versus 3, respectively, Figure 3B). As a result, sequential LysC-trypsin digestion revealed cross-link sites across the whole BSA protein sequence (Figure 3C). Sequential LysC followed by trypsin digestion identified the highest number of unique cross-linked sites, as compared to GluC and chymotrypsin digestions (22 versus 16 and 9, $p < 0.01$ and $p < 0.01$, respectively, Figure 3B). We observed that GluC and chymotrypsin digestions identified 7 and 6 unique cross-link sites, respectively, which were not detected using sequential LysC followed by trypsin digestion (Figure 3C). The improved coverage can be attributed to the accessibility of proteolysis sites and susceptibility to enzymatic endoproteolysis.[61] Comparison of the cross-linked peptide properties generated by the different proteases is shown in Figure S2. Thus, the use of multiple, nonlysine-dependent proteases can expand the coverage and density of DSSO XL-MS protein interaction maps.

Native Chromatin Protein–Protein Interactions Identified Using Cross-Linking Mass Spectrometry in Situ

Having defined improved experimental parameters for large-scale cross-linking mass spectrometry, we applied them to study DSSO-labeled native protein complexes in live human cells. We chose to study protein–protein interactions of human chromatin in situ, given its abundance and prominent functions in the regulation of cell growth and development. We optimized a protocol to label live human HEK293T cells with DSSO (see Experimental Section). Chromatin proteins were extracted from nuclei purified by sucrose sedimentation, followed by sequential salt, detergent, and micrococcal nuclease digestion to release chromatin bound proteins.[64–66] We used SDS-PAGE followed by silver staining to confirm the presence of protein complexes with reduced electrophoretic mobility upon DSSO cross-linking, as compared to mock-treated cells (Figure 4A). Likewise, we used Western immunoblotting to verify the isolation of histone-containing chromatin complexes, as compared to nucleoplasmic or cytoplasmic proteins, as assessed by the detection of histone H3, BRG1, and GAPDH, respectively (Figure 4B).

Analysis of the cross-linked chromatin fraction identified 2252 unique cross-linked sites from three biological replicates, with 624 unique cross-linked sites from at least two out of three biological replicates across four different digestion strategies: (1) trypsin only, (2) LysC followed by trypsin, (3) chymotrypsin, and (4) GluC (Table S1). A comparable number of cross-links were identified with trypsin digestion and with sequential LysC followed by trypsin digestion (384 and 253 unique cross-linked sites respectively), while chymotrypsin and GluC digestion identified 43 and 27 unique cross-linked sites, respectively. Forty cross-link sites identified by chymotrypsin were nonoverlapping with the tryptic or GluC cross-link sites, and 27 cross-link sites identified by GluC were nonoverlapping with the tryptic or chymotryptic cross-link sites. The lower number of cross-linked peptides identified by chymotrypsin and GluC can be attributable to the lower susceptibility of chromatin bound nuclear proteins to proteolytic cleavage by these enzymes and by lower charge state and ionization efficiencies of the produced cross-linked peptides.

However, in consideration of the improved sequence coverage, these results suggest that multiple proteases, and their combination, will be needed for comprehensive protein interaction studies. Three hundred and thirty-seven of the 624 unique cross-linked sites (54%) were interactions within the same protein, either through intra-molecular interaction or between homopolymers, involving 114 unique proteins. Two hundred and eighty-seven (46%) of the 624 unique cross-linked sites were between two different proteins involving 177 unique proteins. One hundred and eighty-six of the 624 interactions are represented in the current BioGRID and CORUM protein–protein interaction databases, which includes self-associating proteins.[67–70] Among the cross-links which passed the 1% FDR of Percolator, 20 and 9 decoy cross-links were identified from trypsin and sequential LysC, trypsin digestion, respectively. No decoy cross-links passed the 1% FDR of Percolator for chymotrypsin and GluC digests. These decoy cross-links have, on average, a lower CSM score than the target cross-links identified (Figure S3A).

Fasci and co-workers recently reported cross-linking mass spectrometry of human cells using DSSO, including the analysis of the TX100 insoluble fraction, which is partially composed of nuclear chromatin.[71] We compared these data with the cross-links identified in our analysis, and observed that 74 of 624 cross-linked sites were shared between the two data sets, in spite of the use of different cell types, extraction buffers and MS fragmentation methods (Figure S3B). Out of these shared identifications, 18 cross-links were between two different proteins and 56 cross-links involved the same protein. These included 33 cross-links involving histone proteins, HMGN1–3, PARP1 and ribo-nucleoproteins. Consistent with the relationship between CSM scores and accuracy, we found that the shared cross-links have substantially higher CSM scores than the average of the entire data set (Figure S3C).

In agreement with the empirically determined CSM score thresholds for identifying high-confidence protein–protein interactions, we observed that all 11 cross-links involving residues in the nucleosome histone core can be mapped onto the high-resolution, atomic-level structure of the human nucleosome,[72] with 11 cross-links conforming to the 26 Å distance constraint based on the length of the DSSO linker (Figure 4C). Likewise, protein–protein interactions involving the core histone proteins recapitulated many known histone–histone interactions, such as those between the core histone H2B and H2A, H4, and H3 histones (Figure 4C). Notably, we observed 60 cross-links involving the histone tails, which were not observed in the isolated nucleosome core structure in vitro,[72] suggesting that they are in fact organized and bound to specific cofactors in situ. In particular, we observed numerous cross-links involving histone H2B K5 and histone H3 K4 and K18, which were also observed in a recent study.[71] The protein–protein interaction network of the core histones (Figure 4D) and histone linker H1 (Figure 4E) included numerous cross-links between the core histone proteins and histone H1, including K4 and K18 of histone H3 to histone H1.[71] Several DNA-binding proteins were also detected as cross-linked to histone proteins, including transcription factor CSRNP2 and zinc finger domain containing ZNF786. The DNA-binding properties of these proteins may suggest an association with histones through proximal binding to DNA, as opposed to the direct binding to histones. Known histone interactors were also identified, such as NCL, a reported interactor with the linker histone H1,[73] and HCFC1, a component of a complex involved in the acetylation of histone

H4.[74] These results suggest that these proteins may have additional structural chromatin functions, possibly in stabilizing or regulating histone linker-dependent higher-order chromatin conformations in situ.

Notably, using stringent identification of cross-links in situ, we observed numerous previously unrecognized interactions. For example, we found numerous cross-links between the HMGN1 and HMGN2 proteins with histone H2B and histone H1.2 (Figure 4D,E). HMGN1 and HMGN2 are known to regulate chromatin configuration,[75] which can involve phosphorylation of histone H2A and histone H3.[76,77] Our findings indicate that HMGN1 can also interact directly with regions surrounding K108 and K120 of H2B, in agreement with recent findings of Fasci and colleagues,[71] and HMGN2 can also interact with K206 of H1.2 in the context of native chromatin in situ. This may provide the sought-after mechanism to explain the structural effects of HMGN1 and HMGN2 on chromatin configuration directly.[78,79] Likewise, we observed new protein–protein interactions involving ADARB1 with histone H2B, and centriolin/CNTRL with histone H3. This suggests that these two proteins may have unanticipated functions on DNA and/or chromatin. In addition, nonhistone proteins were also cross-linked with histone H1, including PHRF1, ZNF786, PDE1C, and SEPT1, suggesting that these proteins may have additional structural chromatin functions, possibly in stabilizing or regulating histone linker-dependent higher-order chromatin conformations in situ (Figure 4E).

Lastly, we found numerous chromatin interactions of proteins involved in DNA damage repair and ribonucleoproteins. For instance, we observed cross-links from K125 of PARP1 to K108 and K116 of histone H2B, which were also observed by Fasci and colleagues,[71] in addition to the 32 intraprotein PARP1 cross-links (Figure 5A). We confirmed that 7 of 32 detected PARP1 cross-links could be mapped onto its high-resolution structure (PDB ID 4dqy),[80] as bound to a DNA double stranded break site (Figure 5B). PARP1 is known to self-associate,[81,82] and our findings provide a direct mechanism of chromatin recruitment, linking PARP1 to histone H2B. This interaction may affect PARP1 self-association and DNA recognition and/or contribute to the chromatin remodeling associated with DNA damage repair.[83–87] Similarly, we found numerous cross-links among the heterogeneous nuclear ribonucleoproteins hnRNPs and chromatin (Figure 5C). This included known interactions between hnRNPA1 and hnRNP2B, and between hnRNPA1 and hnRNPA3, providing specific structural details on their physical proximity.[88,89] In addition, we observed direct interactions between DNA ligase LIG4 and hnRNPU proteins with histone H2B, raising the possibility that chromatin and hnRNP recruitment can regulate repair of DNA double-strand breaks.[90] While these interactions will require dedicated future studies to confirm and establish their functions, these results indicate that native chromatin protein–protein interactions can be identified by using cross-linking mass spectrometry in situ.

CONCLUSIONS AND FUTURE DIRECTIONS

Here, we sought to define improved methods for the detection of cross-linked peptides from protein complexes labeled in situ, and to ensure their accurate identification using a strategy adapted for cross-linking mass spectrometry. We observed that the target-decoy approach, used to establish the confidence of peptide-spectral matching, does not completely eliminate

false identifications of cross-linked peptides. The generalized linear mixture model presented here offers a facile means to control the sensitivity and specificity of cross-linked peptide identifications based on empirically derived CSM score distributions. This approach is scalable and versatile, as additional cross-linked proteins or proteomes can be used in the future to model protein interactions with specific features, such as those containing particular chemical modifications. We also anticipate that scoring functions incorporating fragmentation features of specific cross-linkers as a function of fragmentation and isolation methods, such as the diagnostic sulfoxide fragment ion doublet in DSSO, can be used to improve the sensitivity and specificity of cross-linked identification.[91–93] Likewise, the generation of improved cross-linking reagents, such as isotopically labeled heavy and light cross-linkers, should improve the specificity and sensitivity of cross-link spectral matching. Here, we demonstrated that sample-specific target databases can be used to improve accuracy of spectral matching for the study of specific complexes, and that the granularity and coverage of protein–protein interaction maps can be significantly improved through the use of proteases with diverse sequence specificities. We also found that complex cross-linked mixtures can be effectively analyzed using the relatively fast CID-MS2/HCD-MS2 fragmentation methods. Lastly, our studies of native human chromatin labeled in live cells recapitulated known and revealed new protein interactions of nucleosomes and other chromatin-associated complexes. In all, this approach should facilitate the discovery and definition of protein–protein complexes in vivo, both in health and disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Rosa Viner and David Horn for support of the beta version of XLinkX and Yael David and reviewers for comments on the manuscript. This work was supported by the NIH R01 CA204396, P30 CA008748, the Burroughs Wellcome Fund, the Josie Robertson Investigator Program, the Rita Allen Foundation, the Alex's Lemonade Stand Foundation, St. Baldrick's Arceci Innovation Award, (A.K.), and Agency for Science, Technology and Research, Singapore (Z.S.). A.K. is the Damon Runyon-Richard Lumsden Foundation Clinical Investigator.

REFERENCES

(1). Kuzmanov U; Emili A Protein-protein interaction networks: probing disease mechanisms using model systems. Genome Med 2013, 5 (4), 37. [PubMed: 23635424]

(2). Smits AH; Vermeulen M Characterizing Protein-Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. Trends Biotechnol 2016, 34 (10), 825–34. [PubMed: 26996615]

(3). Huttlin EL; Ting L; Bruckner RJ; Gebreab F; Gygi MP; Szpyt J; Tam S; Zarraga G; Colby G; Baltier K; Dong R; Guarani V; Vaites LP; Ordureau A; Rad R; Erickson BK; Wuhr M; Chick J; Zhai B; Kolippakkam D; Mintseris J; Obar RA; Harris T; Artavanis-Tsakonas S; Sowa ME; De Camilli P; Paulo JA; Harper JW; Gygi SP The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell 2015, 162 (2), 425–40. [PubMed: 26186194]

(4). Huttlin EL; Bruckner RJ; Paulo JA; Cannon JR; Ting L; Baltier K; Colby G; Gebreab F; Gygi MP; Parzen H; Szpyt J; Tam S; Zarraga G; Pontano-Vaites L; Swarup S; White AE; Schweppe DK; Rad R; Erickson BK; Obar RA; Guruharsha KG; Li K; Artavanis-Tsakonas S; Gygi SP; Harper JW Architecture of the human interactome defines protein communities and disease networks. Nature 2017, 545 (7655), 505–509. [PubMed: 28514442]

(5). Hein MY; Hubner NC; Poser I; Cox J; Nagaraj N; Toyoda Y; Gak IA; Weisswange I; Mansfeld J; Buchholz F; Hyman AA; Mann M A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell 2015, 163 (3), 712–23. [PubMed: 26496610]

(6). Havugimana PC; Hart GT; Nepusz T; Yang H; Turinsky AL; Li Z; Wang PI; Boutz DR; Fong V; Phanse S; Babu M; Craig SA; Hu P; Wan C; Vlasblom J; Dar VU; Bezginov A; Clark GW; Wu GC; Wodak SJ; Tillier ER; Paccanaro A; Marcotte EM; Emili A A census of human soluble proteincomplexes. Cell 2012, 150 (5), 1068–81. [PubMed: 22939629]

(7). Budayeva HG; Cristea IM A mass spectrometry view of stable and transient protein interactions. Adv. Exp. Med. Biol 2014, 806, 263–82. [PubMed: 24952186]

(8). Weisz DA; Liu H; Zhang H; Thangapandian S; Tajkhorshid E; Gross ML; Pakrasi HB Mass spectrometry-based cross-linking study shows that the Psb28 protein binds to cytochrome b559 inPhotosystem II. Proc. Natl. Acad. Sci. U. S. A 2017, 114 (9), 2224–2229. [PubMed: 28193857]

(9). Rappsilber J The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J. Struct. Biol 2011, 173 (3), 530–40. [PubMed: 21029779]

(10). Schneider M; Belsom A; Rappsilber J Protein TertiaryStructure by Crosslinking/Mass Spectrometry. Trends Biochem. Sci 2018, 43 (3), 157–169. [PubMed: 29395654]

(11). Muller MQ; Dreiocker F; Ihling CH; 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. [PubMed: 20704385]

(12). Tang X; Bruce JE A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. Mol. BioSyst 2010, 6 (6), 939–47. [PubMed: 20485738]

(13). Kao A; Chiu CL; Vellucci D; Yang Y; Patel VR; Guan S; Randall A; Baldi P; Rychnovsky SD; 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.

(14). Kaake RM; Wang X; Burke A; Yu C; Kandur W; Yang Y; Novtisky EJ; Second T; Duan J; Kao A; Guan S; Vellucci D; Rychnovsky SD; Huang L A new in vivo cross-linking mass spectrometry platform to define protein-protein interactions in living cells. Mol. Cell. Proteomics 2014, 13 (12), 3533–43. [PubMed: 25253489]

(15). Kandur WV; Kao A; Vellucci D; Huang L; Rychnovsky SD Design of CID-cleavable protein cross-linkers: identical mass modifications for simpler sequence analysis. Org. Biomol. Chem 2015, 13 (38), 9793–807. [PubMed: 26269432]

(16). Liu F; Rijkers DT; Post H; Heck AJ Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. Nat. Methods 2015, 12 (12), 1179–84. [PubMed: 26414014]

(17). Liu F; Lossl P; Scheltema R; Viner R; Heck AJR Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. Nat. Commun 2017, 8, 15473. [PubMed: 28524877]

(18). Gotze M; Pettelkau J; Fritzsche R; Ihling CH; Schafer M; Sinz A Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. J. Am. Soc. Mass Spectrom 2015, 26 (1), 83–97. [PubMed: 25261217]

(19). Arlt C; Gotze M; Ihling CH; Hage C; Schafer M; Sinz A Integrated Workflow for Structural Proteomics Studies Based on Cross-Linking/Mass Spectrometry with an MS/MS Cleavable Cross-Linker. Anal. Chem 2016, 88 (16), 7930–7. [PubMed: 27428000]

(20). Mohr JP; Perumalla P; Chavez JD; Eng JK; Bruce JE Mango: A General Tool for Collision Induced Dissociation-Cleavable Cross-Linked Peptide Identification. Anal. Chem 2018, 90 (10), 6028–6034. [PubMed: 29676898]

(21). Yang B; Wu YJ; Zhu M; Fan SB; Lin J; Zhang K; Li S; Chi H; Li YX; Chen HF; Luo SK; Ding YH; Wang LH; Hao Z; Xiu LY; Chen S; Ye K; He SM; Dong MQ Identification of cross-linked peptides from complex samples. Nat. Methods 2012, 9 (9), 904–6. [PubMed: 22772728]

(22). Tan D; Li Q; Zhang MJ; Liu C; Ma C; Zhang P; Ding YH; Fan SB; Tao L; Yang B; Li X; Ma S; Liu J; Feng B; Liu X; Wang HW; He SM; Gao N; Ye K; Dong MQ; Lei X Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states. eLife 2016, 10.7554/eLife.12509.

(23). Zheng C; Yang L; Hoopmann MR; Eng JK; Tang X; Weisbrod CR; Bruce JE Cross-linking measurements of in vivo protein complex topologies. Mol. Cell. Proteomics 2011, 10 (10), M110.006841.

(24). Zhong X; Navare AT; Chavez JD; Eng JK; Schweppe DK; Bruce JE Large-Scale and Targeted Quantitative Cross-Linking MS Using Isotope-Labeled Protein Interaction Reporter (PIR) Cross-Linkers. J. Proteome Res 2017, 16 (2), 720–727. [PubMed: 28152603]

(25). Chavez JD; Weisbrod CR; Zheng C; Eng JK; Bruce JE Protein interactions, post-translational modifications and topologies in human cells. Mol. Cell. Proteomics 2013, 12 (5), 1451–67. [PubMed: 23354917]

(26). Liu F; Lossl P; Rabbitts BM; Balaban RS; Heck AJR The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for coexisting respiratory super-complexes. Mol. Cell. Proteomics 2018, 17 (2), 216–232. [PubMed: 29222160]

(27). Wu X; Chavez JD; Schweppe DK; Zheng C; Weisbrod CR; Eng JK; Murali A; Lee SA; Ramage E; Gallagher LA; Kulasekara HD; Edrozo ME; Kamischke CN; Brittnacher MJ; Miller SI; Singh PK; Manoil C; Bruce JE In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in Acinetobacter baumannii strain AB5075. Nat. Commun 2016, 7, 13414. [PubMed: 27834373]

(28). Weisbrod CR; Chavez JD; Eng JK; Yang L; Zheng C; Bruce JE In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. J. Proteome Res 2013, 12 (4), 1569–79. [PubMed: 23413883]

(29). Wisniewski JR; Zougman A; Nagaraj N; Mann M Universal sample preparation method for proteome analysis. Nat. Methods 2009, 6 (5), 359–62. [PubMed: 19377485]

(30). Wisniewski JR; Zielinska DF; Mann M Comparison of ultrafiltration units for proteomic and N-glycoproteomic analysis by the filter-aided sample preparation method. Anal. Biochem 2011, 410 (2), 307–9. [PubMed: 21144814]

(31). Dhabaria A; Cifani P; Reed C; Steen H; Kentsis A A High-Efficiency Cellular Extraction System for Biological Proteomics. J. Proteome Res 2015, 14 (8), 3403–8. [PubMed: 26153614]

(32). Dhabaria A; Cifani P; Kentsis A Fabrication of Capillary Columns with Integrated Frits for Mass Spectrometry; Protocol Exchange, 2015.

(33). Cifani P; Dhabaria A; Kentsis A Fabrication of Nanoelectrospray Emitters for LC–MS; Kentsis Research Group, 2015.

(34). The UniProt Consortium. UniProt: the universal protein knowledgebase. Nucleic Acids Res 2017, 45 (D1), D158–D169. [PubMed: 27899622]

(35). Mellacheruvu D; Wright Z; Couzens AL; Lambert JP; St-Denis NA; Li T; Miteva YV; Hauri S; Sardiu ME; Low TY; Halim VA; Bagshaw RD; Hubner NC; Al-Hakim A; Bouchard A; Faubert D; Fermin D; Dunham WH; Goudreault M; Lin ZY; Badillo BG; Pawson T; Durocher D; Coulombe B; Aebersold R; Superti-Furga G; Colinge J; Heck AJ; Choi H; Gstaiger M; Mohammed S; Cristea IM; Bennett KL; Washburn MP; Raught B; Ewing RM; Gingras AC; Nesvizhskii AI The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat. Methods 2013, 10 (8), 730–6. [PubMed: 23921808]

(36). Kall L; Canterbury JD; Weston J; Noble WS; MacCoss MJ Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat. Methods 2007, 4 (11), 923–5. [PubMed: 17952086]

(37). Oliveros JC Venny: An Interactive Tool for Comparing Lists with Venn's Diagrams; 2007–2015; http://bioinfogp.cnb.csic.es/tools/venny/index.html.

(38). Lex A; Gehlenborg N; Strobelt H; Vuillemot R; Pfister H UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput. Graph 2014, 20 (12), 1983–92. [PubMed: 26356912]

(39). Combe CW; Fischer L; Rappsilber J xiNET: cross-link network maps with residue resolution. Mol. Cell. Proteomics 2015, 14 (4), 1137–47. [PubMed: 25648531]

(40). Pettersen EF; Goddard TD; Huang CC; Couch GS; Greenblatt DM; Meng EC; Ferrin TE UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem 2004, 25 (13), 1605–12. [PubMed: 15264254]

(41). Kosinski J; von Appen A; Ori A; Karius K; Muller CW; Beck M Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. J. Struct. Biol 2015, 189 (3), 177–83. [PubMed: 25661704]

(42). Shannon P; Markiel A; Ozier O; Baliga NS; Wang JT; Ramage D; Amin N; Schwikowski B; Ideker T Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003, 13 (11), 2498–504. [PubMed: 14597658]

(43). 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. [PubMed: 29793340]

(44). Bhattacharya M; Jain N; Mukhopadhyay S Insights into the mechanism of aggregation and fibril formation from bovine serum albumin. J. Phys. Chem. B 2011, 115 (14), 4195–205. [PubMed: 21417250]

(45). Vetri V; Librizzi F; Leone M; Militello V Thermal aggregation of bovine serum albumin at different pH: comparison with human serum albumin. Eur. Biophys. J 2007, 36 (7), 717–25. [PubMed: 17624524]

(46). Spivak M; Weston J; Bottou L; Kall L; Noble WS Improvements to the percolator algorithm for Peptide identification from shotgun proteomics data sets. J. Proteome Res 2009, 8 (7), 3737–45. [PubMed: 19385687]

(47). Bujacz A Structures of bovine, equine and leporine serum albumin. Acta Crystallogr., Sect. D: Biol. Crystallogr 2012, 68 (10), 1278–89. [PubMed: 22993082]

(48). Shen Y; Tolic N; Xie F; Zhao R; Purvine SO; Schepmoes AA; Moore RJ; Anderson GA; Smith RD Effectiveness of CID, HCD, and ETD with FT MS/MS for degradomic-peptidomic analysis: comparison of peptide identification methods. J. Proteome Res 2011, 10 (9), 3929–43. [PubMed: 21678914]

(49). Good DM; Wirtala M; McAlister GC; Coon JJ Performance characteristics of electron transfer dissociation mass spectrometry. Mol. Cell. Proteomics 2007, 6 (11), 1942–51. [PubMed: 17673454]

(50). Kolbowski L; Mendes ML; Rappsilber J Optimizing the Parameters Governing the Fragmentation of Cross-Linked Peptides in a Tribrid Mass Spectrometer. Anal. Chem 2017, 89 (10), 5311–5318. [PubMed: 28402676]

(51). Stieger CE; Doppler P; Mechtler K Optimized fragmentation improves the identification of peptides cross-linked by MS-cleavable reagents. J. Proteome Res 2019, 18, 1363. [PubMed: 30693776]

(52). Barysz HM; Malmstroem J Development of large-scale cross-linking mass spectrometry. Mol. Cell. Proteomics 2018, 17, 1055. [PubMed: 28389583]

(53). Sticker A; Martens L; Clement L Mass spectrometrists should search for all peptides, but assess only the ones they care about. Nat. Methods 2017, 14 (7), 643–644. [PubMed: 28661493]

(54). Noble WS Mass spectrometrists should search only for peptides they care about. Nat. Methods 2015, 12 (7), 605–8. [PubMed: 26125591]

(55). Noble WS; Keich U Response to “Mass spectrometrists should search for all peptides, but assess only the ones they care about”. Nat. Methods 2017, 14 (7), 644. [PubMed: 28661496]

(56). Knudsen GM; Chalkley RJ The effect of using an inappropriate protein database for proteomic data analysis. PLoS One 2011, 6 (6), No. e20873. [PubMed: 21695130]

(57). Jeong K; Kim S; Bandeira N False discovery rates in spectral identification. BMC Bioinf 2012, 13 (16), S2.

(58). Kumar D; Yadav AK; Dash D Choosing an Optimal Database for Protein Identification from Tandem Mass Spectrometry Data. Methods Mol. Biol 2017, 1549, 17–29. [PubMed: 27975281]

(59). 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.

(60). Dau T; Gupta K; Berger I; Rappsilber J Sequential Digestion with Trypsin and Elastase in Cross-Linking Mass Spectrometry. Anal. Chem 2019, 91, 4472. [PubMed: 30817130]

(61). Giansanti P; Tsiatsiani L; Low TY; Heck AJ Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. Nat. Protoc 2016, 11 (5), 993–1006. [PubMed: 27123950]

(62). Tsiatsiani L; Heck AJ Proteomics beyond trypsin. FEBS J 2015, 282 (14), 2612–26. [PubMed: 25823410]

(63). Swaney DL; Wenger CD; Coon JJ Value of using multiple proteases for large-scale mass spectrometry-based proteomics. J. Proteome Res 2010, 9 (3), 1323–9. [PubMed: 20113005]

(64). Kornberg RD; LaPointe JW; Lorch Y Preparation of nucleosomes and chromatin. Methods Enzymol 1989, 170, 3–14. [PubMed: 2770543]

(65). Hewish DR; Burgoyne LA Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem. Biophys. Res. Commun 1973, 52 (2), 504–10. [PubMed: 4711166]

(66). Zaret K Micrococcal nuclease analysis of chromatin structure. Curr. Protoc Mol. Biol 2005, 10.1002/0471142727.mb2101s69, Chapter 21, Unit 21 1.

(67). Stark C; Breitkreutz BJ; Reguly T; Boucher L; Breitkreutz A; Tyers M BioGRID: a general repository for interaction datasets. Nucleic Acids Res 2006, 34, D535–9. [PubMed: 16381927]

(68). Chatr-Aryamontri A; Oughtred R; Boucher L; Rust J; Chang C; Kolas NK; O'Donnell L; Oster S; Theesfeld C; Sellam A; Stark C; Breitkreutz BJ; Dolinski K; Tyers M The BioGRID interaction database: 2017 update. Nucleic Acids Res 2017, 45 (D1), D369–D379. [PubMed: 27980099]

(69). Ruepp A; Brauner B; Dunger-Kaltenbach I; Frishman G; Montrone C; Stransky M; Waegele B; Schmidt T; Doudieu ON; Stumpflen V; Mewes HW CORUM: the comprehensive resource of mammalian protein complexes. Nucleic Acids Res 2007, 36, D646–D650. [PubMed: 17965090]

(70). Ruepp A; Waegele B; Lechner M; Brauner B; Dunger-Kaltenbach I; Fobo G; Frishman G; Montrone C; Mewes HW CORUM: the comprehensive resource of mammalian protein complexes–2009. Nucleic Acids Res 2010, 38, D497–D501. [PubMed: 19884131]

(71). Fasci D; van Ingen H; Scheltema RA; Heck AJR Histone Interaction Landscapes Visualized by Crosslinking Mass Spectrometry in Intact Cell Nuclei. Mol. Cell. Proteomics 2018, 17 (10), 2018–2033. [PubMed: 30021884]

(72). Tachiwana H; Osakabe A; Shiga T; Miya Y; Kimura H; Kagawa W; Kurumizaka H Structures of human nucleosomes containing major histone H3 variants. Acta Crystallogr., Sect. D: Biol. Crystallogr 2011, 67 (Pt 6), 578–583. [PubMed: 21636898]

(73). Kalashnikova AA; Rogge RA; Hansen JC Linker histone H1 and protein-protein interactions. Biochim. Biophys. Acta, Gene Regul. Mech 2016, 1859 (3), 455–61.

(74). Cai Y; Jin J; Swanson SK; Cole MD; Choi SH; Florens L; Washburn MP; Conaway JW; Conaway RC Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. J. Biol. Chem 2010, 285 (7), 4268–72. [PubMed: 20018852]

(75). Murphy KJ; Cutter AR; Fang H; Postnikov YV; Bustin M; Hayes JJ HMGN1 and 2 remodel core and linker histone tail domains within chromatin. Nucleic Acids Res 2017, 45 (17), 9917–9930. [PubMed: 28973435]

(76). Lim JH; Catez F; Birger Y; West KL; Prymakowska-Bosak M; Postnikov YV; Bustin M Chromosomal protein HMGN1 modulates histone H3 phosphorylation. Mol. Cell 2004, 15 (4), 573–84. [PubMed: 15327773]

(77). Postnikov YV; Belova GI; Lim JH; Bustin M Chromosomal protein HMGN1 modulates the phosphorylation of serine 1 in histone H2A. Biochemistry 2006, 45 (50), 15092–9. [PubMed: 17154547]

(78). Shimahara H; Hirano T; Ohya K; Matsuta S; Seeram SS; Tate S Nucleosome structural changes induced by binding of nonhistone chromosomal proteins HMGN1 and HMGN2. FEBS Open Bio 2013, 3, 184–91.

(79). Postnikov YV; Bustin M Functional interplay between histone H1 and HMG proteins in chromatin. Biochim. Biophys. Acta, Gene Regul. Mech 2016, 1859 (3), 462–7.

(80). Langelier MF; Planck JL; Roy S; Pascal JM Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1. Science 2012, 336 (6082), 728–32. [PubMed: 22582261]

(81). Wacker DA; Ruhl DD; Balagamwala EH; Hope KM; Zhang T; Kraus WL The DNA binding and catalytic domains of poly(ADP-ribose) polymerase 1 cooperate in the regulation of chromatin structure and transcription. Mol. Cell. Biol 2007, 27 (21), 7475–85. [PubMed: 17785446]

(82). Bonfiglio JJ; Fontana P; Zhang Q; Colby T; Gibbs-Seymour I; Atanassov I; Bartlett E; Zaja R; Ahel I; Matic I Serine ADP-Ribosylation Depends on HPF1. Mol. Cell 2017, 65 (5), 932–940. [PubMed: 28190768]

(83). Poirier GG; de Murcia G; Jongstra-Bilen J; Niedergang C; Mandel P Poly(ADP-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure. Proc. Natl. Acad. Sci. U. S. A 1982, 79 (11), 3423–7. [PubMed: 6808510]

(84). D'Amours D; Desnoyers S; D'Silva I; Poirier GG Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. Biochem. J 1999, 342 (2), 249–68. [PubMed: 10455009]

(85). Messner S; Altmeyer M; Zhao H; Pozivil A; Roschitzki B; Gehrig P; Rutishauser D; Huang D; Caflisch A; Hottiger MO PARP1 ADP-ribosylates lysine residues of the core histone tails. Nucleic Acids Res 2010, 38 (19), 6350–62. [PubMed: 20525793]

(86). Huambachano O; Herrera F; Rancourt A; Satoh MS Double-stranded DNA binding domain of poly(ADP-ribose) polymerase-1 and molecular insight into the regulation of its activity. J. Biol. Chem 2011, 286 (9), 7149–60. [PubMed: 21183686]

(87). Ray Chaudhuri A; Nussenzweig A The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. Nat. Rev. Mol. Cell Biol 2017, 18 (10), 610–621. [PubMed: 28676700]

(88). Roy R; Durie D; Li H; Liu BQ; Skehel JM; Mauri F; Cuorvo LV; Barbareschi M; Guo L; Holcik M; Seckl MJ; Pardo OE hnRNPA1 couples nuclear export and translation of specific mRNAs downstream of FGF-2/S6K2 signalling. Nucleic Acids Res 2014, 42 (20), 12483–97. [PubMed: 25324306]

(89). Jurica MS; Licklider LJ; Gygi SR; Grigorieff N; Moore MJ Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA 2002, 8 (4), 426–39. [PubMed: 11991638]

(90). Polo SE; Blackford AN; Chapman JR; Baskcomb L; Gravel S; Rusch A; Thomas A; Blundred R; Smith P; Kzhyshkowska J; Dobner T; Taylor AM; Turnell AS; Stewart GS; Grand RJ; Jackson SP Regulation of DNA-end resection by hnRNPU-like proteins promotes DNA double-strand break signaling and repair. Mol. Cell 2012, 45 (4), 505–16. [PubMed: 22365830]

(91). Ivanov MV; Levitsky LI; Lobas AA; Panic T; Laskay UA; Mitulovic G; Schmid R; Pridatchenko ML; Tsybin YO; Gorshkov MV Empirical multidimensional space for scoring peptide spectrum matches in shotgun proteomics. J. Proteome Res 2014, 13 (4), 1911–20. [PubMed: 24571493]

(92). Frank AM A ranking-based scoring function for peptide-spectrum matches. J. Proteome Res 2009, 8 (5), 2241–52. [PubMed: 19231891]

(93). Fischer L; Rappsilber J Quirks of Error Estimation in Cross-Linking/Mass Spectrometry. Anal. Chem 2017, 89 (7), 3829–3833. [PubMed: 28267312]

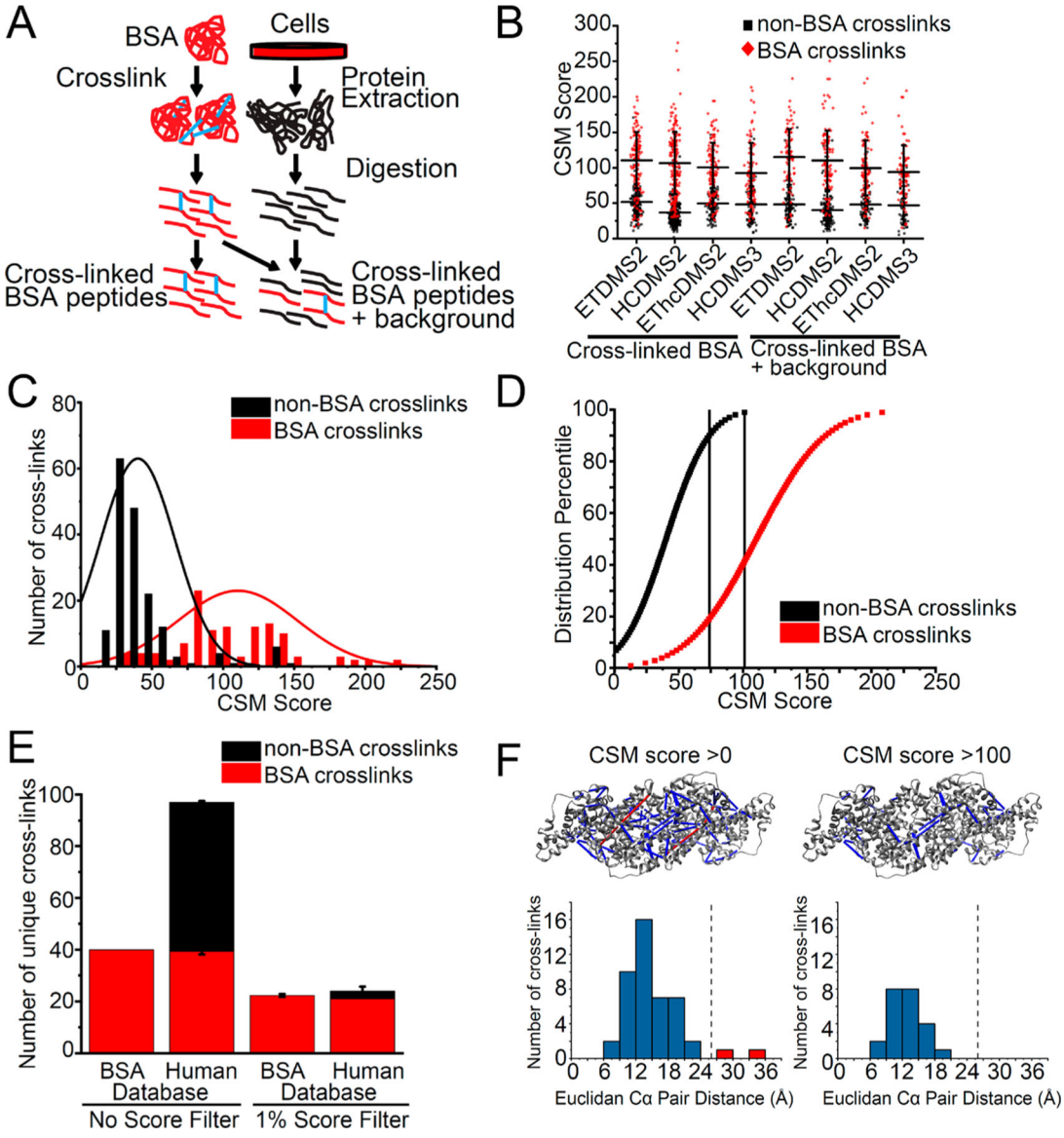


Figure 1.

Calibration of cross-link peptide-spectra matching (CSM) scores to control the sensitivity and specificity of large-scale XL-MS. (A) Workflow for preparation of cross-linked BSA peptides and cross-linked BSA peptides spiked into non-cross-linked proteome background peptides. (B) High number of BSA cross-linked peptides (red) and non-BSA cross-linked peptides (black) identified across fragmentation methods and in both cross-linked BSA and cross-linked BSA spiked into background. (C) Fitted Gaussian distributions of cross-link spectra matching score of non-BSA cross-links and BSA cross-links for cross-linked BSA only sample acquired with CID-MS2/HCD-MS2 fragmentation. (D) Percentile plot of Gaussian distributions of CSM scores with 1% and 10% score filter to eliminate 99% and 90% of non-BSA cross-links. (E) Comparison of number of cross-links identified when no score filter is applied and when 1% score filter is applied. (F) Cross-links violating physical distance constraint of cross-linker (red) are eliminated when 1% score filter is applied when

mapping cross-links to crystal structure (PDB ID 4f5s). Dotted line indicates maximum Cα–Cα distance of 26 Å for DSSO.

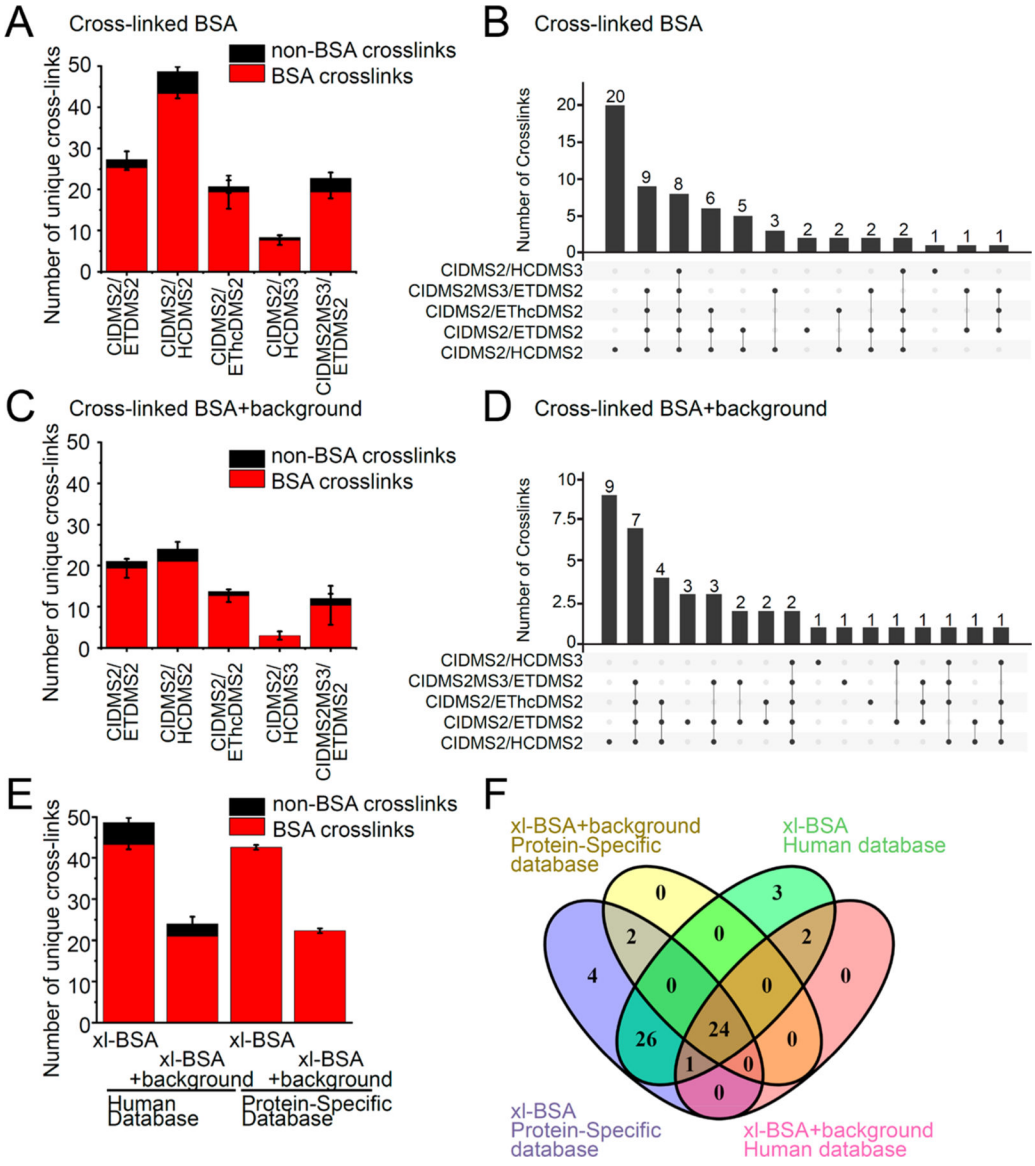


Figure 2.

Improved cross-linked peptide identification using CID-MS2/HCD-MS2 fragmentation. Comparison of 4 different fragmentation methods: CID-MS2/HCD-MS2, CID-MS2/ETD-MS2, CID-MS2/EThcD-MS2, CID-MS2/HCD-MS3, and CID-MS2-MS3/ETD-MS2 and databases: BSA protein specific database or human proteome database with BSA for increasing cross-link peptide identification. (A) Comparison between 5 fragmentation methods for average number of cross-link identifications of cross-linked BSA sample. BSA cross-links colored red and non-BSA cross-links colored black. Error bars represent standard deviation from three biological replicates. (B) Comparison between 5 fragmentation methods for overlap of all unique cross-link identifications. (C) Comparison between 5 fragmentation methods for average number of cross-link identifications of cross-linked BSA peptides spiked into human proteome background. (D) Comparison between 5 fragmentation methods for overlap of all unique cross-link identifications of cross-linked BSA sample

spiked into human proteome background. (E) Comparison between 2 databases for average number of cross-link identifications of cross-linked BSA peptides and cross-linked BSA peptides spiked into human proteome background. (F) Comparison between 2 databases for overlap of all unique cross-link identifications of cross-linked BSA peptides and cross-linked BSA peptides spiked into human proteome background.

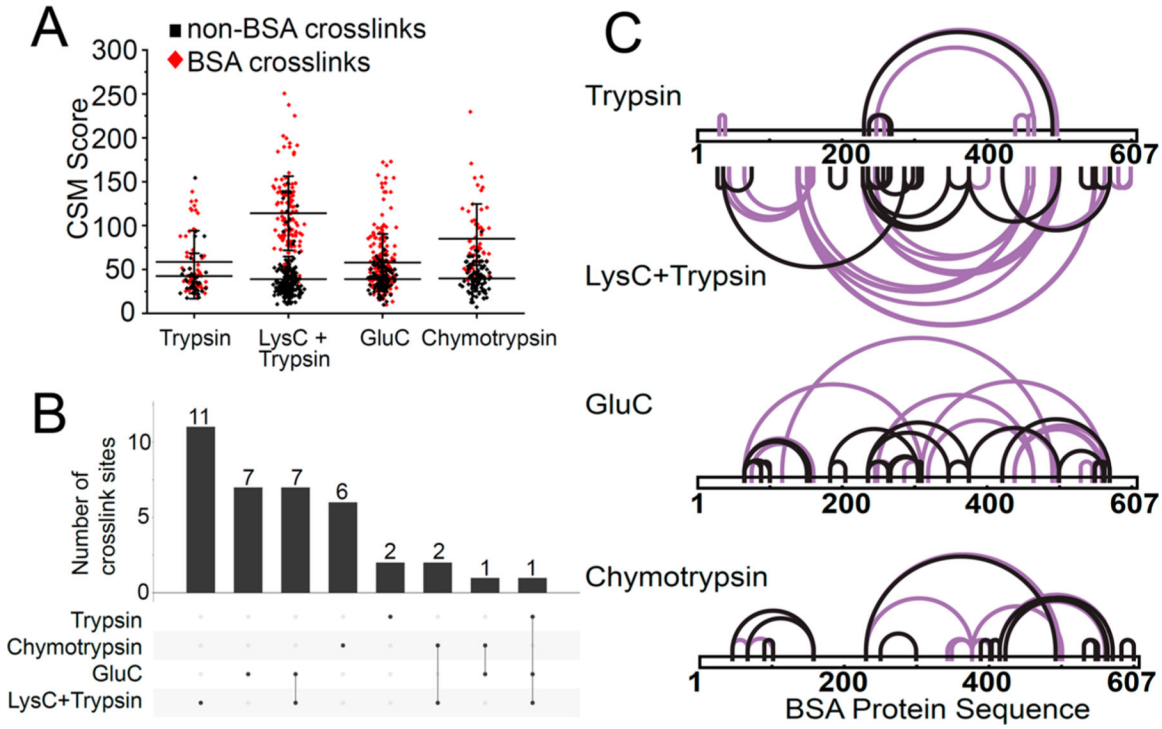


Figure 3.

Multiple proteases to expand the coverage and density of DSSO XL-MS protein interaction maps. (A) Distribution of CSM score for identified cross-linked peptides from cross-linked BSA digested with trypsin only or sequential LysC and trypsin or GluC only or chymotrypsin only proteases. Cross-linked BSA peptides were spiked into non-cross-linked proteome background digested with the corresponding protease. (B) Overlap of cross-linked lysine sites identified from trypsin only or sequential LysC and trypsin or GluC only or chymotrypsin only protease digestion. (C) Visualization of BSA cross-link sites identified from trypsin only, LysC and trypsin, GluC only or chymotrypsin only protease digestion, as analyzed using CID-MS2/HCD-MS2 fragmentation. Cross-links with a CSM score above the 1% score filter are highlighted in black.

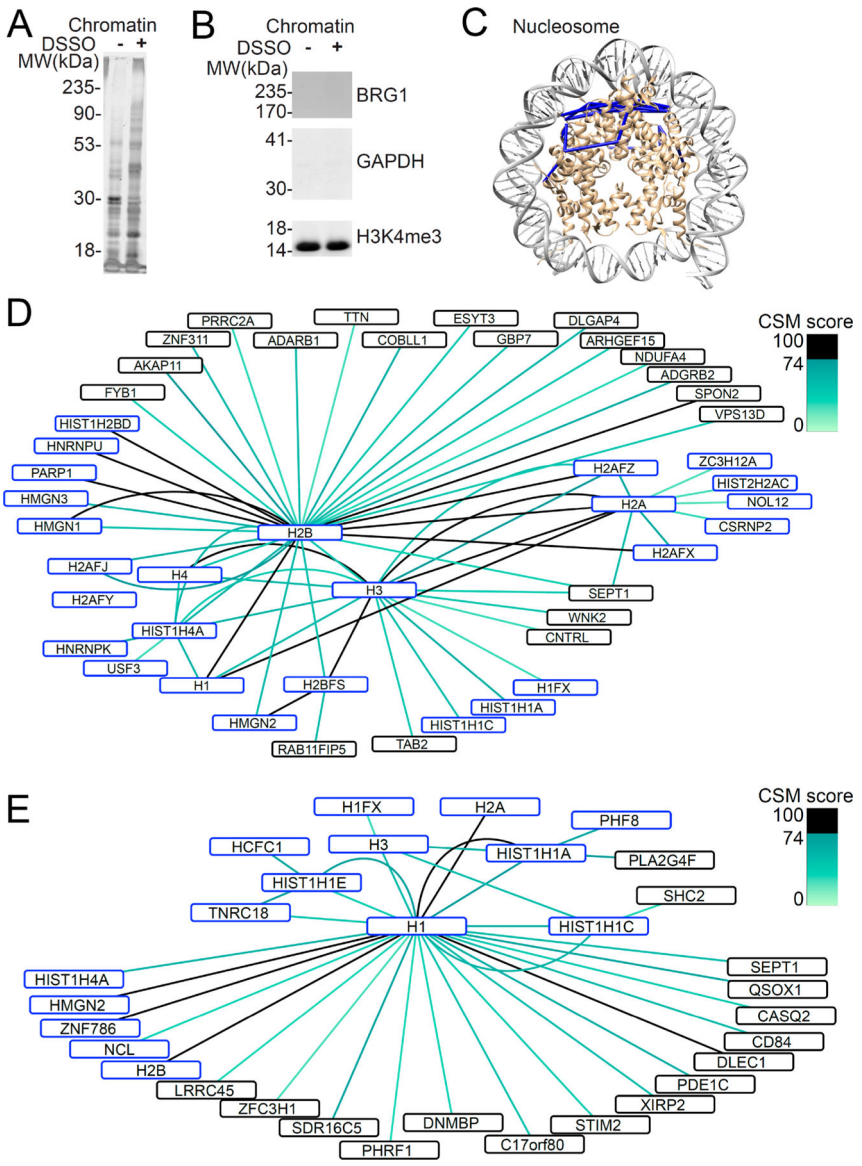


Figure 4.

Native chromatin protein–protein interactions identified using cross-linking mass spectrometry in situ. (A) Purified chromatin bound proteins are cross-linked by cross-linking in situ. (B) Enrichment of histone protein (H3K4me3) with absence of nuclear protein (BRG1) and cytoplasmic protein (GAPDH) from chromatin bound proteins by Western blot. (C) Cross-links identified map onto known high resolution nucleosome structure (PDB ID 3av1). (D) Map of protein–protein interactions between core histone proteins (H2A, H2B, H3, H4) colored according to CSM score. (E) Map of protein–protein interactions involving linker histone H1. Proteins known to bind DNA are marked in blue.

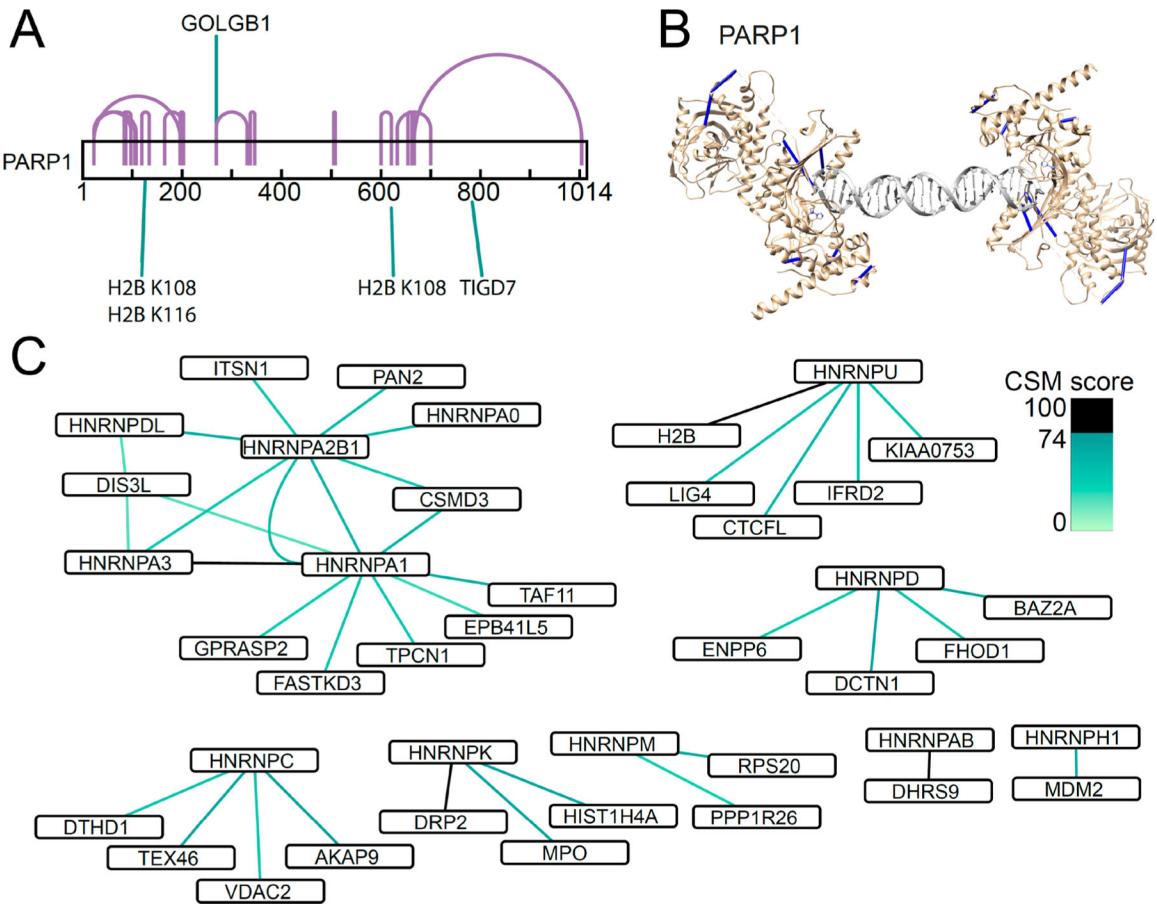


Figure 5.

Chromatin interactions of proteins involved in DNA damage repair and ribonucleoproteins. (A) DNA repair protein PARP1 interacts with Histone H2B. (B) PARP1 cross-links map onto known high resolution structure of PARP1 bound to DNA double stranded break (PDB ID 4dqy). (C) Protein–protein interactions of RNA-binding heterogeneous nuclear ribonucleoproteins (hnRNPs).