Structural Analysis of Protein Complexes by Cross-Linking and Mass-Spectrometry

Moriya Slavin and Nir Kalisman

Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, Jerusalem, Israel

Abstract

Cross-linking and mass-spectrometry is used more and more for the structural analysis of large proteins and protein complexes. Although essentially a low-resolution method, it avoids the main drawbacks of established structural techniques. Particularly, it is largely insensitive to the inherent flexibility of the studied complexes, and is applied under native conditions. It is also applicable to nearly every structural system. Therefore, cross-linking and mass-spectrometry is the method of choice for elucidating the general architecture of protein complexes. Advances in instrumentation, techniques, and software now allow every lab that is working with proteins to apply the approach without much difficulty. The most delicate step in the workflow, the mass-spectrometry measurement, can be done in most facilities that are performing standard proteomics. We detail here a step-by-step protocol of how to successfully apply the approach in collaboration with the mass-spectrometry facility in your institution.

Keywords: Structural biology, Molecular machines, Protein architecture

1 Introduction

Structural studies of large protein complexes are very challenging. Established techniques such as crystallography or cryo-electron microscopy are often not applicable to such complexes because of their inherent flexibility. Additionally, large complexes might be unstable and break into heterogeneous mixtures under the preparation conditions of these approaches. In recent years, the method of cross-linking and mass-spectrometry (abbreviated as XL-MS or CX-MS) has emerged as an attractive alternative that largely avoid such issues [1]. In XL-MS, the protein complex is incubated with a short bifunctional cross-linking reagent under native conditions (Figure 1). The cross-linking reaction creates stable covalent links between protein side-chains that are structurally close within the context of the intact complex. The proteins are then denatured and digested into peptides, some of which are still cross-linked in pairs. Through the 'peptide sequencing' capability of the mass-spectrometer, the sequences of two cross-linked peptides can be found, thus identifying the exact pair of residues that underwent cross-linking. This 'residue resolution' is one of the main advantages of XL-MS. A typical XL-MS application on a protein complex will result in hundreds of cross-link identifications, which in turn are

converted into connectivity maps and distance constraints. This is a very rich structural resource that elucidates the architecture of the protein complex.

Cross-linking and mass-spectrometry was a key methodology in recent structural works that covered such diverse systems as chaperones, transcription or photosynthesis [2-5]. In these works the cross-links were mainly used to report on the interactions between different protein subunits within complex. In many cases the cross-link data were combined with other structural information sources. For example, excellent synergy occurs between cross-links and electron densities obtained by cryo-electron microscopy. The electron density gives the global envelope and shape of the particle, while the cross-links constrain locally the possible arrangements of the subunits within that envelope.

The advances in instrumentation and software now allow any lab to easily apply XL-MS on its protein complex of interest. The wide applicability, inherent simplicity, and low cost of the XL-MS workflow make it the approach of choice for the initial structural probing of most systems. The most specialized step in the workflow is the mass-spectrometry measurement, which requires an expensive mass-spectrometry system. Fortunately, these days many research institutions have a mass-spectrometry facility to which samples are submitted for proteomics analysis. Since both XL-MS and standard proteomics are run on the same instruments, we assume that most labs could measure their cross-linked samples at their local MS facility. Accordingly, we wrote this protocol to be used by labs that have the standard molecular biology equipment but not their in-house mass-spectrometer.

This protocol describes cross-linking with two specific cross-linking reagents – BS^3 (bis-sulfosuccinimidyl suberate) and DSS (disuccinimidyl suberate). These are very popular regents that were used in a large fraction of the recent publications utilizing XL-MS. Their popularity arises from several factors. First, they specifically cross-link only primary amines, which on proteins means the side-chains of lysine residues and the N-termini. This specificity greatly simplifies the analysis of the cross-link data. Second, they are highly reactive at the physiological pH of 7.5. Finally, they strike a good balance of being short enough to convey meaningful structural information, and yet sufficiently long to ensure cross-linking. We therefore highly recommend their use if compatible with the specific system.

2 Materials

2.1 Cross-linking and Digestion

2.1.1 The DSS or BS^3 cross-linking reagents: Both DSS or BS^3 are moisture-sensitive. Their powders should be stored in a desiccator at 4°C. DSS can be dissolved in DMSO to a concentration of 100mM and then divided into aliquots and stored at -80°C for many months. BS^3 solution should be prepared fresh immediately before use. BS^3 can be dissolved in PBS or HEPES buffers to a concentration of 100mM.

2.1.2 Buffers compatible with cross-linking: Make sure that your sample is in a buffer that is compatible with the cross-linking reaction. If buffer is not compatible, then buffer exchange must precede the cross-linking. For both BS3 and DSS, the pH of the buffer must be above 7.0 and below 8.5. The buffer components must not contain primary amines that would react with those regents. HEPES (50mM) or PBS (1x) are good buffers, while TRIS is not because of its primary amine moiety.

2.1.3 1M Ammonium bicarbonate solution for quenching.

2.1.4 Acetone – HPLC grade.

2.1.5 Urea buffer: 8M urea solution in DDW. Urea buffer should be prepare fresh before the MS preparation.

2.1.6 Urea/DTT: 8M urea buffer with 10mM DTT.

2.1.7 Iodoacetamide: Prepare a 0.5M solution of IAA in urea buffer. IAA is light-sensitive and moisture sensitive. Stored desiccated at 4°C. Prepare solution just before use and perform reaction in the dark.

2.1.8 Digestion buffer: 25mM Tris-HCl pH 8, 10%ACN.

2.1.9 Sequencing grade trypsin: Vendors commonly sell lyophilized trypsin. 20ug of lyophilized trypsin powder should be reconstituted in 40ul of 50mM acetic acid, divided into aliquots of 1-2ug trypsin in each, and stored in -80°C.

2.1.10 Acidifying: 5% Trifluoroacetic acid (TFA) in DDW.

2.2 Desalting on Stage Tips

2.2.1 C18 resin: Empore Solid Phase Extraction Octadecyl (C18) 47mm diameter disks (from 3M).

2.2.2 200ul pipette tips, 2ml and 1.5ml collection tubes.

2.2.3 All solvents in this section should be at least HPLC grade and preferably MS grade.

2.2.4 Wetting solution: 50% water, 50% ACN, 0.1% TFA.

2.2.5 Washing solution: 0.1% TFA in water.

2.2.6 Elution solution: 25% water, 75% acetonitrile, 0.1% formic acid.

2.2.7 MS reconstitution solution: 97% water, 3% acetonitrile, 0.1% formic acid.

2.3 Size Exclusion Chromatography (SEC)

2.3.1 SEC buffer: 70% water, 30% acetonitrile, 0.1% TFA – all HPLC grade.

2.3.2 Sep-Pak C18 cartridges (Note 1 ; Waters).

2.3.3 Superdex Peptide PC 3.2/30 column (GE Systems) and HPLC system.

3 Methods

3.1 Cross-linking the Protein Sample

3.1.1 Cross-linking With the BS^3 Reagent

3.1.1.1 Make sure that your sample is in a buffer that is compatible with the cross-linking reaction.

3.1.1.2 Prepare a concentrated solution of 100mM BS^3 in buffer and add from it to the protein sample for the desired final concentration of BS^3. Mix well. It is recommended to use the minimal BS^3 concentration that would still provide sufficient inter-subunit crosslinking. Previous studies [2,5] have shown the optimal concentration of cross-linker to be 1-3mM for systems of purified protein complexes. For very concentrated samples (such as lysates) you can go up to 10mM to ensure the cross-linking reagent is not dwindled by the sample.

3.1.1.3 Incubate the protein sample with the cross-linker for 45min at 30°C or for 90min on ice.

3.1.1.4 Add 30mM ammonium bicarbonate to quench any unreacted BS^3 in the sample and incubate for additional 15min.

3.1.1.5 At this stage native conditions do not matter anymore as cross-linking has completed. This is a good stopping point and sample can be frozen and processed at a later time.

3.1.2 Cross-linking with the DSS Reagent

3.1.2.1 The cross-linking protocol for DSS is essentially the same as that for BS^3. Both DSS and BS^3 have the same cross-linking reactivity toward primary amines and the same 8-carbon spacer. Unlike BS^3, DSS is membrane permeable and particularly suited for in-vivo cross-linking. Because DSS is not water-soluble, it must be dissolved in DMSO before being added to the sample (Note 2).

3.1.2.2 Add the DSS solution in DMSO to the protein sample for the desired final concentration of DSS. Mix well. The addition of DSS will cause the solution to turn opaque because of DMSO micellization.

3.2 Sample Preparation for Mass-Spectrometry

3.2.1 Acetone Precipitation of Proteins (Note 3)

3.2.1.1 Place your sample in an acetone-resistant tube (e.g. 1.5ml Eppendorff tubes). Estimate your sample volume and add at least 5-fold of acetone to it. Mix well.

3.2.1.2 Cool the sample in -80°C for one hour.

3.2.1.3 Centrifuge for 10min at 14,000g, preferably under maximal cooling.

3.2.1.4 Discard the supernatant without disturbing the pellet. Do not dry the pellet or it would be very difficult to re-solubilize. Small amounts of residual acetone will evaporate in subsequent steps.

3.2.2 Protein Denaturation, Reduction, Alkylation, and Digestion

3.2.2.1 Add 20µL of urea/DTT solution to the protein pellet. Pipette vigorously until all the pellet is re-suspended (Note 4). Incubate at 30°C for 30min. This step denatures the proteins and breaks any disulfide bonds between cysteine residues.

3.2.2.2 Add 2.2ul of Iodoacetamide to final concentration of 50mM and incubate at 30°C for 30min in the dark. This step modifies the side-chains of cysteine residues so that they cannot reform disulfide bonds.

3.2.2.3 Dilute the sample by 10-fold with 220ul of digestion buffer. Add the trypsin and mix well. Use a trypsin-to-protein mass ratio of 1:50-100. Incubate the trypsin digestion for 12-18 hours at 37°C while shaking (600 rpm).

3.2.2.4 Stop the trypsin digestion by acidifying the sample. Adding 5% TFA solution to 1/10 the sample volume will reduce the pH to 4.

3.2.2.5 Place the sample in a speed-vac for 20min to evaporate the acetonitrile.

3.3 Desalting on Stage-Tips

Salts and urea must be removed from the sample before it is injected to the mass-spectrometer. To that aim, the peptides are loaded onto C18 resin and bind to it through hydrophobic interactions. The salts do not bind and are washed away. The peptides are then eluted from the C18 with acetonitrile. Here we detail an inexpensive and efficient way to desalt the sample by using tips and a centrifuge [6].

3.3.1 Stage Tips Preparation

3.3.1.1 With a blunt bore needle cut a plug with a diameter of 1mm through a disk of Empore C18 resin (Fig. 2a). Squeeze the resin plug into the bottom of a standard 200ul pipette tip. More plugs can be added to increase the peptide loading capacity of the tip. The loading capacity of a tip with three plugs is about 7ug of peptides.

3.3.1.2 Punch a hole in the cap of 2ml tube and place the tip through the hole so that it is held half way. Close the tube with the cap and place in a centrifuge. The tip is now kept suspended above the bottom of the tube. Any solvent on top of the resin will be eventually forced through it into the by centrifugation.

3.3.2 Desalting

3.3.2.1 Add 50µL of wetting solution onto the tip and centrifuge at 1,500g for 2 min.

3.3.2.2 Add 100µL of washing solution and centrifuge at 1,500g for 2 min. Repeat. Empty the 2ml collection tube.

3.3.2.3 Add the acidified peptide digest on top of the resin and centrifuge at 1,500g for 2min.

3.3.2.4 Add 100µL of washing solution and centrifuge at 1,500g for 2 min. Repeat. In the last wash make sure all the solution on top of the resin washed through. If some wash solution is left, continue to centrifuge until cleared (Note 5).

3.3.2.5 Cut off the caps from new 1.5ml tubes. Transfer the tips and punctured caps onto the new tubes.

3.3.2.6 Add 20ul of elution solution and centrifuge at 1,000g for 1 min. Repeat.

3.3.2.7 Remove the tip and cap and place the tube with the eluted peptides in a speed-vac for 11-13 minutes until dryness.

3.3.2.8 Reconstitute the peptides by adding the MS reconstitution solution to the tube and mix well. Peptides are now ready for mass-spectrometry measurement. The exact volume of reconstitution depends on the amount of peptides and the settings of the chromatography system that is coupled to the mass-spectrometer. The amount of peptides can be measured by A_{280} absorbance in a Nano-Drop instrument. Typically you will want to inject to the mass-spectrometer about 0.5ug of your peptides in a volume of about 3ul.

3.4 Enrichment of Cross-linked Peptides

Linear peptide rather than cross-linked peptides are the great majority in the trypsin digest. Therefore, enrichment of the digest for cross-linked peptides will lead to better results. While enrichment is optional, it may increase the number of identified cross-links by more than 50%. Here we present an enrichment method that is based on size exclusion chromatography (SEC) [7]. It utilizes the molecular weight of the cross-linked peptides, which is on average twice as that of linear peptides.

3.4.1 Buffer Exchange by Sep-Pak C18 Cartridge

3.4.1.1 Add 200µL of wetting solution to the cartridge. Use a syringe to applying air pressure to the top of the cartridge and force the solution through the resin to a waste tube.

3.4.1.2 Add 200µL of washing solution and force it through the resin. Repeat.

3.4.1.3 Add the acidified peptide digest on top of the resin and force it through the resin.

3.4.1.4 Add 200µL of washing solution and force it through the resin. Repeat.

3.4.1.5 Prepare new 1.5ml tubes to collect the eluted peptides in the next step.

3.4.1.6 Add 150ul of elution solution and force it through the resin to the new 1.5ml tube. Repeat.

3.4.1.7 Place the tube in a speed-vac until dryness.

3.4.1.8 Reconstitute the peptides in 20ul of SEC buffer.

3.4.2 Size Exclusion Chromatography (Note 6)

3.4.2.1 Load the peptides in SEC buffer onto the Superdex Peptide column at a flow rate of 50ul/min. Collect fractions every 100ul. A typical chromatogram is shown in Fig. 2b. Fractions collected at flow volumes of 1.3-1.6ml are the richest in cross-linked peptides. Elution volumes of 1.6-1.8ml also contain some cross-linked peptides.

3.4.2.2 Dry completely the relevant fractions in a speed-vac. Mix peptides with MS reconstitution solution as discussed in subsection 3.3.2.8.

3.5 Measurement in the Mass-Spectrometer

3.5.1 Mass-spectrometers coupled to reverse-phase liquid chromatography (LC-MS) are now available for proteomics analysis in many research institutions. The same mass-spectrometers can also measure cross-linked samples. However, even after enrichment the cross-linked peptides are still a minor component of the total peptide content of the sample (the majority being linear peptides). Because most mass-spectrometers will first measure the more abundant peptides, the standard proteomics setting is slightly changed to increase the chances of measuring cross-linked peptides. Therefore, inform your mass-spectrometry facility of the following subsections when submitting the sample.

3.5.2 Cross-linking will lead to samples that are more complex than the original protein content. Longer gradients are therefore beneficial. Use a 45 minute gradient for cross-linked samples of 2-3 proteins. Use a 90-120 minute gradient for cross-linked samples of larger protein complexes.

3.5.3 Set data-dependent triggering for MS/MS fragmentation that selects for ions with charge +3 or higher. For more complex samples with >5 proteins select only for ions with charge +4 or higher. These settings are required because cross-linked peptides typically have a charge of +4 or higher, and never less than +3.

3.5.4 From our experience, the optimal fragmentation energy for cross-linked peptides during MS/MS is the same as that for linear peptides.

3.5.5 We measure our cross-linked samples on a Q-Exactive Plus mass-spectrometer with the following settings: Buffer A – water with 0.1% formic acid, Buffer B – acetonitrile with 0.1% formic acid. Gradient rising linearly from 0% buffer B to 45% buffer B over 90 minutes, then rising to 80% buffer B over 5 minutes ; Full MS resolution 70,000 ; MS1 AGC target 1e6 ; MS1 Maximum IT 200ms ; Scan range 250 to 1800 ; dd-MS/MS resolution 35,000 ; MS/MS AGC target 2e5 ; MS2 Maximum IT 300ms; Loop count Top 12 ;Isolation window 1.5 ; Fixed first mass 130 ; MS2 Minimum AGC target 800 ; Charge exclusion: unassigned,1,2,3,8,>8 ; Peptide match off ; Exclude isotope on ; Dynamic exclusion 45sec.

3.6 Analysis of the Mass-Spectrometry Data (Note 7)

3.6.1 Convert the mass-spectrometer output files from RAW to MGF format. A good conversion tool is the Proteome Discoverer software, which is installed in MS facilities that use Thermo mass-spectrometers.

3.6.2 Download the software package - Find_XL - at http://biolchem.huji.ac.il/nirka/software.html. Unzip the package and find a text document with full instructions for its use in the top folder. The package also comes with example data files from a cross-linking experiment on RNA Polymerase II. Use this example to test the installation and get hands-on experience.

3.6.3 The analysis requires two inputs: (i) The MGF files of the MS data. Multiple files can be used and consolidated into a single non-redundant cross-link list. (ii) The sequences of the proteins in the studied complex. Follow the instructions on how to set the paths to the input files.

3.6.4 The output is a text file containing a list of identified cross-links. The cross-links are sorted by decreasing confidence scores. Therefore, only the top of the list should be considered. The question of where to cut the list depends on the False Positive Rate (FPR) that you expect from the data. To determine the FPR, the analysis is concurrently running also the protein sequences in reverse. Consequently, the output list will contain entries with the 'REV' annotation that are clearly false positives. The list should be cut so that the ratio of of 'REV' entries divided by the total number of entries above the cut is the desired FPR.

3.6.5 An excellent visualization tool for the connectivity map implied by the cross-links (Fig. 2c) is Circos [8], which can be accessed at: http://cx-circos.net/ .

4 Notes

4.1 Up to 200ug of peptides can be loaded onto this C18 device.

4.2 The freezing temperature of DMSO is 19°C. If it is used for cross-linking on ice, the DSS-DMSO solution must be largely diluted by the sample.

4.3 Acetone precipitation will remove some detergents (such as Triton) but not others (such as SDS). Since detergents are mostly incompatible with mass-spectrometry, they should be removed by other means or avoided altogether.

4.4 If the pellet is very large or hard to solubilize you can add more urea/DTT. However, this might lead to very large volume of the sample in subsequent steps and should be avoided if possible.

4.5 This is a good stopping point. Peptides can be stored on the C18 resin for many months before being eluted. Add some washing solution on top of the resin to prevent it from drying and store at 4°C.

4.6 SEC on narrow columns of such small volumes is challenging. Avoid dead volumes by using narrow tubing and removal of unnecessary devices in the flow path.

4.7 Here we describe how to use our analysis software- 'Find_XL'. Other software options are 'Xi' from the Rappsilber lab at: http://rappsilberlab.org/rappsilber-laboratory-home-page/tools/ , or 'Stravox' from the Götze lab at: http://www.stavrox.com/ .

5 References

1. Walzthoeni T, Leitner A, Stengel F, Aebersold R (2013) Mass Spectrometry Supported Determination of Protein Complex Structure. Current Opinion in Structural Biology 23: 252-260.
2. Kalisman N, Adams, CM, Levitt M (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. PNAS 109:2884-2489.
3. Robinson P, Trnka, M, Bushnell D, Davis R, Mattei P, Burlingame A and Kornberg R (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. Cell, 166:1411-1422.
4. Liu H, Zhang H, Niedzwiedzki D, Prado M, He G, Gross M, Blankenship R (2013) Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. Science 342:1104-1107.
5. Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills JC, Nilges M, Cramer P, Rappsilber J (2010) Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J 29:717-726.

6. Rappsilber J, Mann M and Ishihama Y (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Protocols, 2:1896-1906.
7. Leitner A, Reischl R, Walzthoeni T, Herzog F, Bohn S, Förster F, Aebersold R (2012) Expanding the Chemical Cross-Linking Toolbox by the Use of Multiple Proteases and Enrichment by Size Exclusion Chromatography. Mol Cell Proteomics 11(3):M111.014126.
8. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an Information Aesthetic for Comparative Genomics. Genome Res 19:1639-1645

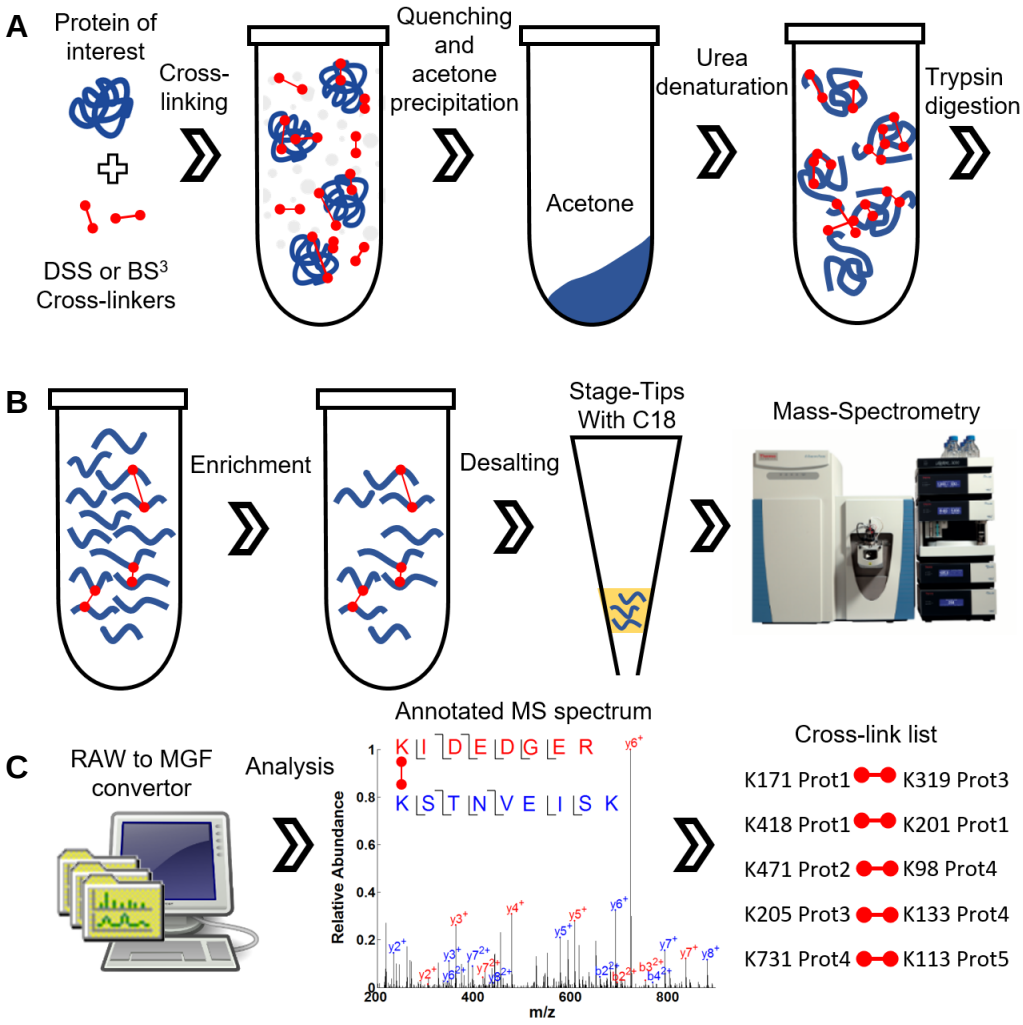

Fig. 1 The XL-MS workflow. (**a**) The protein complex is incubated with a short bi-functional cross-linking reagent under native conditions. Subsequent steps extract the cross-linked protein, denature it, and digest it with trypsin. (**b**) The resulting peptide digest is enriched for the cross-linked peptides by size exclusion chromatography. The sample is desalted and measured in the mass-spectrometer. (**c**) The mass-spectrometer output files are processed and searched for spectra that report on cross-linked peptide pairs. An example of such spectrum shows the overlapping fragmentation series from two peptides, which indicates that they were cross-linked into a single ion. The result of the analysis is a list of cross-links that fully details all the pairs of residues (protein names and residue numbers) that underwent cross-linking.

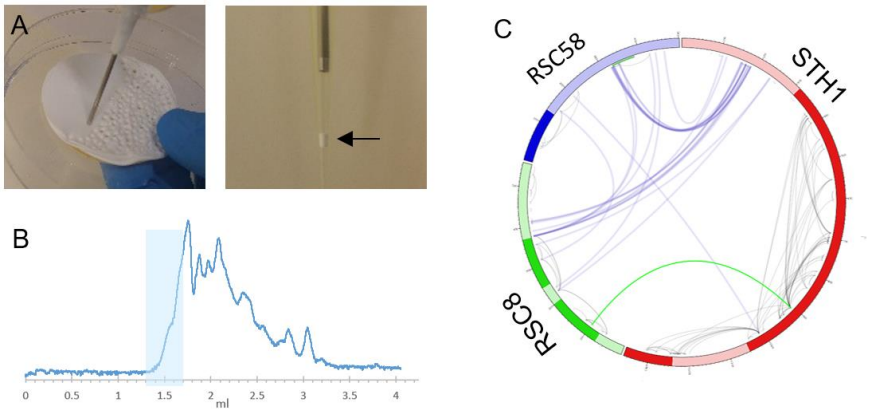

Fig. 2 Steps in the workflow. (**a**) Preparation of the desalting stage-tips. A small plug of C18 resin is cut by a blunt bore needle (left) and then pushed to the bottom of a 200ul pipette tip (arrow on right). (**b**) Chromatogram of SEC enrichment of cross-linked peptides from a large protein complex. The shaded area mark the fractions in which most of the cross-links are later identified. (**c**) CX-Circos is a powerful visualization tool for cross-link data. Here the cross-links (arcs) between three proteins in a complex are plotted. Specific domains in the proteins and their cross-links can be colored differently.