

# Systems structural biology measurements by in vivo cross-linking with mass spectrometry

Juan D. Chavez<sup>1,2</sup>, Jared P. Mohr<sup>1,2</sup>, Martin Mathay<sup>1</sup>, Xuefei Zhong<sup>1</sup>, Andrew Keller<sup>1</sup> and James E. Bruce<sup>1\*</sup> 

**This protocol describes a workflow for utilizing large-scale cross-linking with mass spectrometry (XL-MS) to make systems-level structural biology measurements in complex biological samples, including cells, isolated organelles, and tissue samples. XL-MS is a structural biology technique that provides information on the molecular structure of proteins and protein complexes using chemical probes that report the proximity of probe-reactive amino acids within proteins, typically lysine residues. Information gained through XL-MS studies is often complementary to more traditional methods, such as X-ray crystallography, nuclear magnetic resonance, and cryo-electron microscopy. The use of MS-cleavable cross-linkers, including protein interaction reporter (PIR) technologies, enables XL-MS studies on protein structures and interactions in extremely complex biological samples, including intact living cells. PIR cross-linkers are designed to contain chemical bonds at specific locations within the cross-linker molecule that can be selectively cleaved by collision-induced dissociation or UV light. When broken, these bonds release the intact peptides that were cross-linked, as well as a reporter ion. Conservation of mass dictates that the sum of the two released peptide masses and the reporter mass equals the measured precursor mass. This relationship is used to identify cross-linked peptide pairs. Release of the individual peptides permits accurate measurement of their masses and independent amino acid sequence determination by tandem MS, allowing the use of standard proteomics search engines such as Comet for peptide sequence assignment, greatly simplifying data analysis of cross-linked peptide pairs. Search results are processed with XLinkProphet for validation and can be uploaded into XlinkDB for interaction network and structural analysis.**

## Introduction

XL-MS is a powerful technique for studying the structures and interactions of proteins<sup>1,2</sup>. Chemical cross-linking reagents with a variety of chemical features have been described<sup>3</sup>. MS-cleavable cross-linkers contain one or more labile bonds that are selectively cleavable within the mass spectrometer upon input of energy sufficient to cause dissociation of labile bonds engineered within the cross-linker<sup>4</sup>. Selective release of the cross-linked peptides allows determination of the individual peptide masses and independent sequence determination by MS<sup>3</sup> analysis. This overcomes the inherent combinatorial problem, known as the  $n^2$  problem<sup>5</sup>, which arises due to the need to consider all possible pairs of peptides when assigning cross-linked peptide sequences to mass spectra obtained with non-cleavable cross-linkers. This issue limits the use of non-cleavable cross-linkers to studying relatively low-complexity systems<sup>6,7</sup>. The use of MS-cleavable cross-linkers overcomes this issue and has allowed large-scale interactome measurements from extremely complex systems, including virus particles, cell lysates<sup>8,9</sup>, isolated organelles<sup>10–12</sup>, intact cells<sup>13–19</sup>, and tissue samples<sup>20</sup>. Indeed, the rapid growth of the XL-MS field is reflected by the recent availability of multiple protocols that detail specific approaches toward XL-MS experiments taken by different laboratories<sup>21–23</sup>.

Since 2005, our laboratory has focused on the development of PIR technologies<sup>24,25</sup> as a means of proteome-wide cross-linking carried out in biological systems of extreme complexity as described in the examples above. The PIR approach utilizes chemical cross-linkers with specific engineered properties, including selectively cleavable bonds (bonds that can undergo collision-induced dissociation (CID) or that are photocleavable) to release the intact cross-linked peptides and a reporter ion for mass spectrometric analysis, and affinity tags (biotin) for enrichment of low-abundance cross-linked products from complex protein and enzymatic digest samples. PIR cross-linkers are peptide-based molecules, which afford the benefits of modular synthesis on the basis of the widespread

<sup>1</sup>Department of Genome Sciences, University of Washington, Seattle, WA, USA. <sup>2</sup>These authors contributed equally: Juan D. Chavez, Jared P. Mohr.  
\*e-mail: [jimbruce@uw.edu](mailto:jimbruce@uw.edu)

availability of solid-phase peptide synthesis (SPPS) chemistry and biological compatibility, as these cross-linkers readily penetrate cellular membranes.

An expanding array of chemical cross-linker options are available to researchers, including MS-cleavable molecules (see Sinz<sup>4</sup> for recent review), and photocleavable<sup>26,27</sup> and electrochemically cleavable options<sup>28,29</sup>. Several MS-cleavable cross-linkers are now commercially available, including DSSO<sup>30</sup> and DBSU<sup>31</sup>, available from Thermo Scientific; and CBDPS<sup>32</sup>, available from Creative Molecules. Researchers planning XL-MS experiments should carefully consider the different options available to them, including the molecular features and properties of cross-linker molecules, to ensure compatibility with their biological systems and experimental design. Regardless of the cross-linker used, information from these experiments generates distance restraints between the surface-exposed proximal amino acids that are reactive with the chemical cross-linker used (typically Lys). One differentiating feature of chemical cross-linker molecules is their spacer arm length, which is ultimately used to generate an upper bound on distance restraints for structural analysis. For example, DSSO has a reported spacer arm length of 10.3 Å, whereas the peptide-based PIR cross-linker with the amino acid sequence succinate-Asp-Pro-Lys(Pro-Asp-succinate)-Lys(biotin)-Gly, known as BDP, has a spacer arm length of 29.3 Å<sup>20</sup>. Despite these differences, results from large-scale cross-linking experiments indicate that the distributions of observed distances from identified cross-links from either cross-linker are similar when mapped to structural models obtained from the Protein Data Bank (PDB)<sup>10,11</sup>. Importantly, this observation highlights that molecular features of proteins themselves, such as backbone and side-chain flexibility, residue reactivity, and accessibility, can be more important determinants of which cross-linked sites are ultimately observed than the spacer arm of the cross-linker used. Although PIR cross-linkers are relatively large molecules as compared with smaller cross-linker molecules such as DSSO or DBSU, results from large-scale *in vivo* cross-linking efforts do not indicate any detrimental effects due to the size. Cross-linked proteins from all sub-cellular locations, including the plasma membrane, cytosol, nucleus, and mitochondria, are represented in the results, indicating that PIR cross-linkers are able to penetrate biological membranes and react with proteins within the cellular environment. Electron and fluorescence microscopy results corroborate these findings<sup>17,33</sup>.

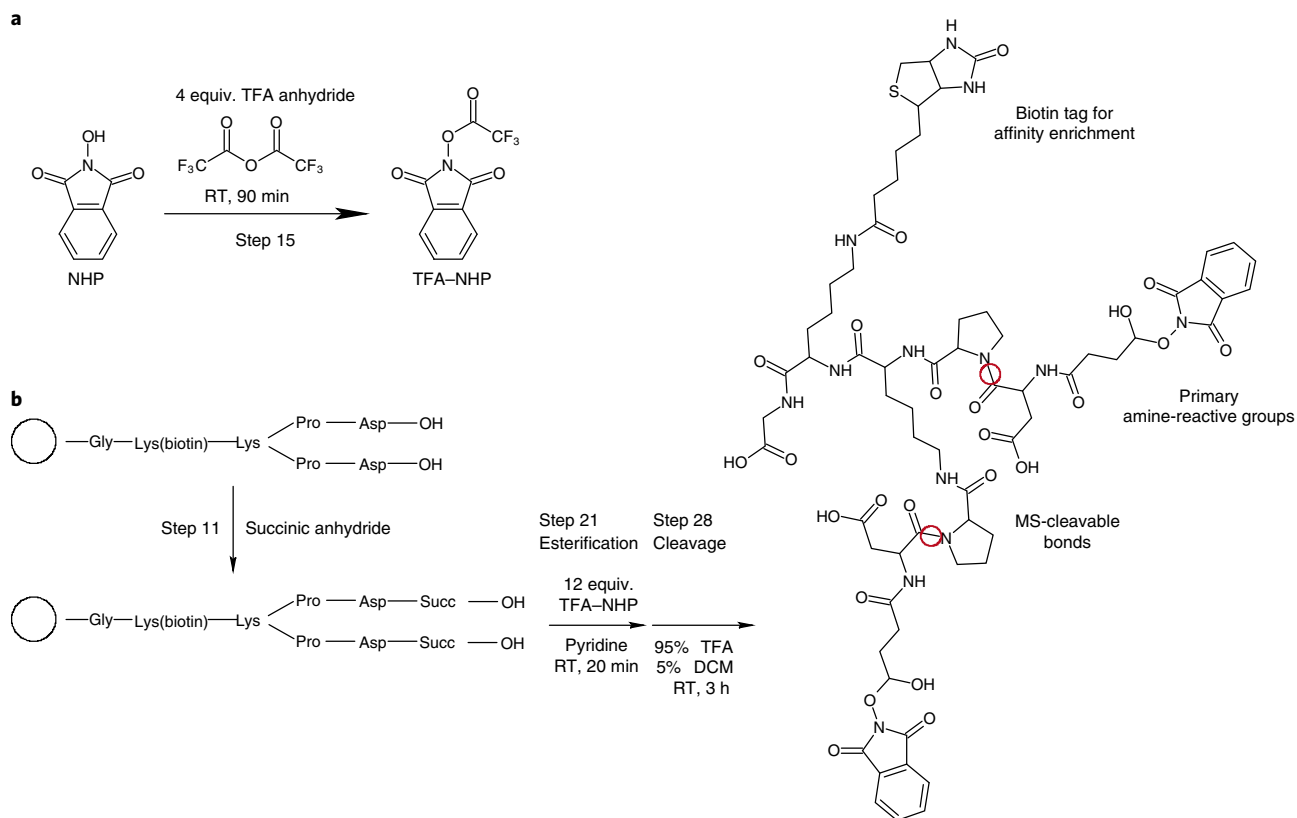
Chemical cross-linking provides information on the spatial proximity of reactive residues on the surfaces of proteins. The generated distance restraints can be used to guide molecular modeling and docking experiments for proteins and protein complexes. The information generated from XL-MS experiments is often complementary to that obtained from other structural techniques such as X-ray crystallography, electron microscopy, and NMR spectroscopy<sup>2,34</sup>. Intermolecular cross-links provide information on protein–protein interactions, including the identity of interacting partners and regions near the interaction interface. Intramolecular cross-links provide information on the structure and conformation of proteins. For *in vivo* cross-linking studies, in which protein structures and interactions are highly dynamic and in their native molecularly crowded environment, XL-MS is able to provide a snapshot from an ensemble of protein conformations and interactions that exist in cells during cross-linking. Here, we provide a protocol for applying large-scale cross-linking using the PIR approach.

### Development of the protocol

The development of the protocol described herein has been a continual process since the initial application of PIR cross-linking in 2005 (ref. <sup>24</sup>). Although numerous improvements and variations have been made in the years since the original inception of the PIR concept, the core foundation of the method remains the same. Fundamentally, the approach depends on the implemented design features of the PIR cross-linker, including amine-reactive functional groups, selectively labile bonds (either collisionally activated or photo-activated cleavage), and an affinity tag for enrichment (biotin) (Fig. 1). Selective cleavage of the labile bonds allows mass spectrometric measurements to define a mass relationship shown in Eq. 1.

$$\begin{aligned}\text{MASS PRECURSOR} &= \text{MASS PEPTIDE1} \\ &+ \text{MASS PEPTIDE2} + \text{MASS REPORTER}\end{aligned}\tag{1}$$

The affinity tag allows for enrichment of cross-linked peptide pairs, which are formed in relatively low abundance, from a complex sample such as an enzymatic digest of a complete proteome. Since the original implementation, recent improvements, including addition of strong cation exchange (SCX) fractionation<sup>35</sup>, the detection of mass relationships in real-time<sup>16</sup>, and detection and identification of the mass relationship and peptide sequences from chimeric spectra<sup>36</sup>, have enabled



**Fig. 1 | Synthesis and molecular features of PIR cross-linker. a**, Synthesis of *N*-hydroxyphthalimide trifluoroacetate ester (TFA-NHP) from NHP and TFA anhydride (Steps 13–16). **b**, Synthesis and molecular features of the PIR cross-linker BDP-NHP. The final product is a bifunctional amine-reactive cross-linker that contains a biotin affinity tag and two MS labile Asp-Pro peptide bonds, denoted with red circles (Steps 1–39). RT, room temperature.

routine detection and identification of thousands of non-redundant cross-linked peptide pairs from *in vivo* cross-linking experiments. Currently, XLinkDB<sup>37–39</sup> represents the largest database of identified cross-linked peptides and contains 19,798 non-redundant cross-linked peptide pairs, the majority of which were acquired with *in vivo* the cross-linking methods described in this protocol. Beyond identification, the incorporation of isotope labels—either metabolically or within the cross-linker molecule—allows for extension of the method for quantitative measurements<sup>40–43</sup>.

### Application of the method

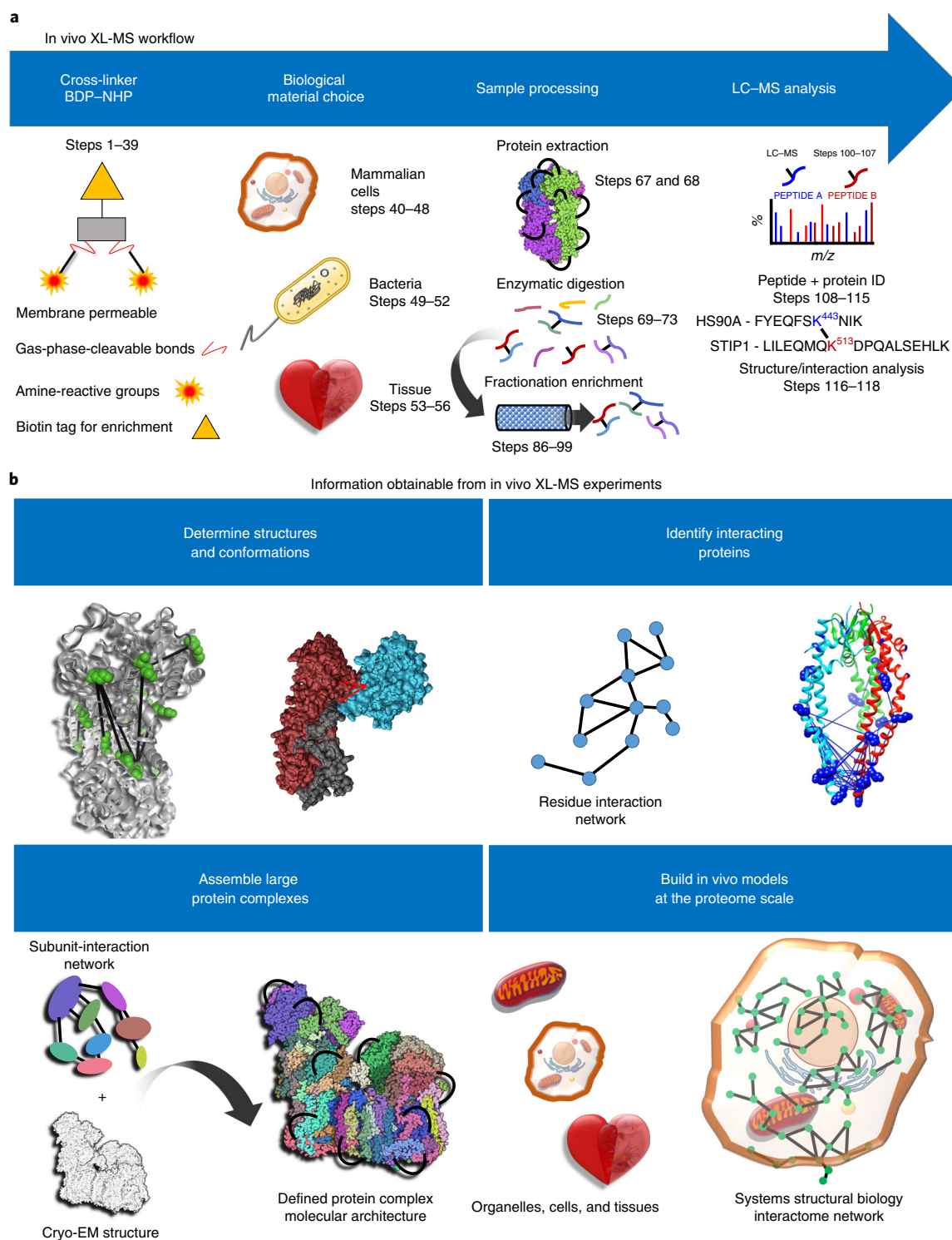
The described method has been successfully applied to large-scale protein structural and interaction measurements in a number of biological systems, including viruses<sup>44–46</sup>, cultured bacterial<sup>15,16,19,47</sup> and mammalian cells<sup>17,42,41</sup>, and tissue samples<sup>20</sup> (Fig. 2). The method requires a range of skills in sample preparation, MS analysis, and data analysis, as with most cross-linking experiments. Our experience with collaborator laboratories indicates that successful PIR experiments can be carried out once a researcher has learned the details provided in this protocol<sup>45,46,48</sup>. Here, we describe the protocol for application of large-scale *in vivo* cross-linking to cultured bacterial and mammalian cell lines, as well as to murine tissue samples.

### Experimental design

It is important to consider the multiple levels of information that are generated by large-scale XL-MS when planning experiments. Fundamentally, identification of cross-linked peptide pairs will provide information on proximal amino acid residues within protein molecules. The identities of interacting proteins are obtained by identifying cross-linked peptide sequences specific to particular protein molecules.

### Cross-linker production

PIR cross-linkers are peptide-based cross-linkers that contain lysine-reactive esters, MS-cleavable bonds, and a biotin affinity tag<sup>24,25</sup>. Being that they are peptide-based, they are readily synthesized



**Fig. 2 | Experimental overview of in vivo XL-MS.** **a**, General workflow includes applying cross-linker (BDP-NHP) to a biological material of choice, which can include mammalian cells, bacterial cells, or animal tissue samples. This is followed by protein extraction, enzymatic digestion, fractionation, and enrichment of cross-linked peptide pairs. Cross-linked peptide pairs are identified by LC-MS analysis. **b**, Types of information resulting from in vivo XL-MS experiments include protein structural and conformational information. XL-MS data can be integrated with data from cryo-EM or other structural biology techniques to assemble structural models for large protein complexes. Identification of interacting proteins can be used to assemble proteome-scale interactome maps for cross-linked systems, including organelles, cells, and tissues.

using standard SPPS protocols as described previously<sup>16,24</sup> and outlined in Steps 1–11. They are also available from suppliers offering custom peptide synthesis (i.e., AnaSpec or others). Variations of PIR molecules have been described, including MS-cleavable versions utilizing the Rink functional group<sup>49</sup>

or the Asp-Pro bond<sup>17</sup> as labile bonds that cleave under relatively low input of CID energy. Photocleavable versions of PIR cross-linkers have also been developed and successfully used for large-scale *in vivo* cross-linking experiments<sup>27</sup>. The final synthesis step for PIR cross-linkers is the formation of reactive ester groups on the cross-linker by reaction with a trifluoroacetate ester, typically *N*-hydroxysuccinimide (NHS) or *N*-hydroxyphthalimide (NHP), and is described below in Steps 13–39 and illustrated in Fig. 1a. After esterification, the cross-linker is cleaved from the solid-phase resin using a solution of 95% (vol/vol) trifluoroacetic (TFA) acid and 5% (vol/vol) dichloromethane (DCM), which also deprotects the Boc-protected Asp side chains.

### Generation of biological sample

Regardless of the type of biological sample to be studied—cultured mammalian cells, bacterial cells, or animal tissue—sample sizes that can yield ~5–10 mg of total protein, as measured by either a Bradford or bicinchoninic acid (BCA) protein assay as in Step 72, is required. Cross-linked peptide pairs are formed at low frequency in the sample, and this starting amount ensures enough material to form cross-linked peptides at a detectable level (femtomole to picomole sensitivity levels) with modern LC–MS instrumentation.

### Cross-linking reaction and sample preparation

The cross-linking reaction must be carried out in an amine-free, aqueous buffer. Sodium phosphate is typically used, but other buffer systems, including HEPES and MOPS, have also been used successfully. As the reactive ester groups on the cross-linker also react with water to form hydrolyzed products, it is important to perform the reaction on a relatively concentrated suspension of cells or tissue to help mitigate hydrolysis. In contrast to cross-linking in purified systems, in which over-cross-linking can be a concern, *in vivo* cross-linking requires relatively large concentrations of cross-linker to ensure formation of cross-linked peptide pairs. This is because, in a cellular sample containing 5 mg of total protein suspended in 0.5 mL of buffer solution, the free amine concentration from Lys residues alone is estimated to be on the order of 10 mM. This is near the solubility limit of BDP–NHP in aqueous solution and only an estimated 1% of the cross-linker will actually react with both of its arms to link Lys residues in proteins because the remainder is hydrolyzed. Thus, it is important to keep in mind that the cross-linker is most often the limiting reagent with *in vivo* experiments and that, in many such cases, quenching of residual cross-linker is not necessary. One potential concern is the possible distortion a protein's structure by the cross-linker molecule. A recent study investigating this concern found that, under *in vitro* reaction conditions, local structural distortions were observed only at high cross-linker concentrations, and even then the overall protein fold was maintained<sup>50</sup>. Under *in vivo* cross-linking conditions, the cross-linker concentration is much lower and unlikely to cause substantial distortion to protein structures. This is supported by the fact that the vast majority (~95%) of cross-linked peptide pairs identified from large-scale *in vivo* cross-linking studies are in excellent agreement with structural models available in the PDB<sup>41</sup>. Furthermore, cross-links that appear to violate their expected distance constraints when mapped to a particular protein structural model, so called 'over-length' cross-links, have in fact been found to contain useful information on different conformations of proteins<sup>51</sup>. Over-length cross-links were particularly important for revealing conformational changes induced in heat shock protein 90 (Hsp90) upon interaction with small-molecule inhibitors<sup>41</sup>.

Control samples can be useful for troubleshooting purposes and ensuring successful experiments. To eliminate potential sources of contamination and ensure specific signal from cross-linked peptide pairs, a non-cross-linked negative-control sample can be generated by adding an equivalent volume of (dimethyl sulfoxide) DMSO instead of PIR cross-linker to a duplicate of the cross-linked biological sample. This sample can be taken through all steps of the protocol, followed by comparison of the resulting MS signal with that of the cross-linked sample. It should not yield any confident cross-linked peptide pair identifications. A positive-control sample can be generated by reacting 1 mL of BSA at 1 mg/mL in 170 mM Na<sub>2</sub>HPO<sub>4</sub> with 1 mM PIR cross-linker and following the protocol from Step 58.

Cell lysis with 8 M urea, followed by sonication, is generally sufficient for complete solubilization of the sample and extraction of protein. Digestion of proteins with trypsin cleaves the peptide bond at the C-terminal side of Arg and Lys residues. Lys residues that have reacted with the cross-linker are no longer positively charged and will result in missed tryptic cleavage sites. Therefore, digestion of cross-linked proteins results in cross-linked peptides containing an internal Lys residue, which has been modified by the cross-linker. This fact can be exploited to improve sensitivity for cross-linked



peptides during the database search by altering the Comet (or other search engine) parameters file (Step 114) to only consider peptide sequences with internal modified Lys residues.

### Liquid chromatography and MS analysis

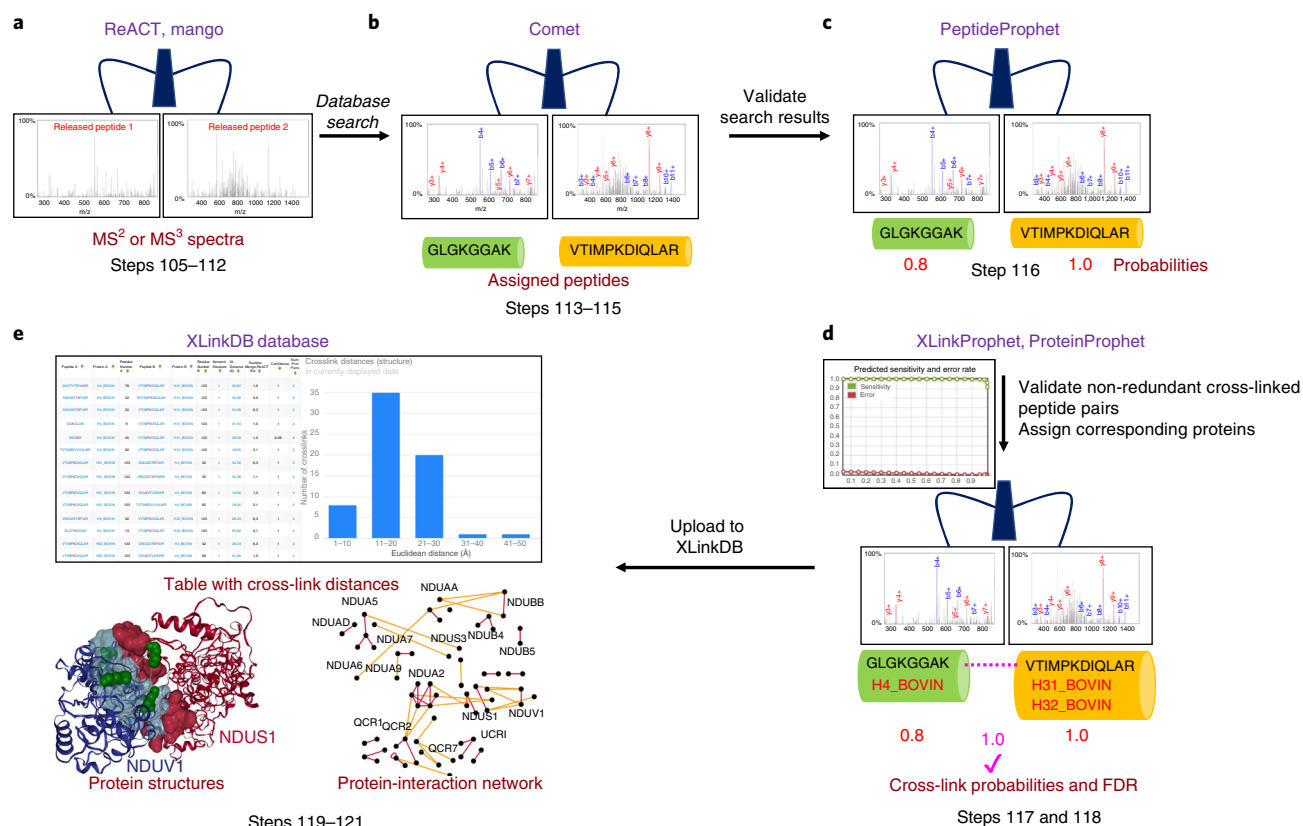
As cross-linked peptide pairs are larger than normal tryptic peptides, the use of LC columns packed with C4 or C8, as opposed to the standard C18, has been shown to be beneficial for reversed-phase chromatographic separation of cross-linked peptides<sup>52</sup>. In addition, the charge states of cross-linked peptide pairs tend to be higher than those of non-cross-linked peptides, so it is beneficial to perform MS2 only on precursor ions with a charge state of 4+ or greater<sup>16,36</sup>. The MS labile bonds cleave at lower CID energy (20 V normalized collision energy) than the peptide backbone, which can be utilized to identify the PIR mass relationship (Eq. 1). Upon cleavage of the MS labile bonds, the individual peptides, each now carrying a mass modification to the cross-linked residue, are released, allowing their masses to be determined. The mass relationship in Eq. 1 can be exploited in real time to have the mass spectrometer sequence only ions that fulfill this criteria, improving sensitivity for cross-linked peptide pairs in a method termed ReACT<sup>16</sup>. However, exploitation of mass relationship information to make downstream decisions, such as triggering MS3, during LC–MS acquisition is not a standard feature available on most commercial instrumentation. To overcome this limitation, we have developed a method to identify both the mass relationship and peptide fragmentation information from chimeric spectra in a post-processing analysis step termed mango<sup>36</sup>. As the mango-based method is able to be executed on any commercial MS instrument with a high resolving power and MS2 capability and will therefore be applicable to the widest range of potential users, this protocol will focus on this method.

### Data processing

A major benefit of using cleavable cross-linkers over non-cleavable versions is the ability to measure the individual masses of the two peptides that were linked. This reduces problems resulting from having to consider all possible combinations of peptides being linked together. The combinatorial expansion of search space hinders the use of non-cleavable cross-linkers in extremely complex samples, such as cells, in which many thousands of protein sequences need to be considered. Use of cleavable cross-linkers also allows standard open-source proteomics search engines such as Comet<sup>53</sup> to be used for searching the resulting MS data. It is important to consider false-discovery rate (FDR) estimation when processing XL-MS data. Owing to the nature of XL-MS data and the multiple levels of information that are gathered—cross-linked peptide sequences, cross-linked residues, and cross-linked proteins—the FDR or FDR estimation is more complicated than that for traditional proteomics approaches<sup>54</sup>. Traditionally, the FDR can be estimated by performing a target/decoy search in which one of three possibilities arise for cross-linked peptide pairs: (i) both peptides are assigned target sequences, (ii) one peptide is assigned a target sequence and the other a decoy sequence, and (iii) both peptides are assigned decoy sequences. Only the first possibility results in a positive identification, as long as both peptides receive a score, such as an expectation value (*e*-value), beyond some acceptable level. Because the situation can arise that one peptide in the cross-linked peptide pair receives a very good score and the other a not-so-good score, the conservative approach has been to assign a score to the peptide pair that is only as good as the worse score of the two peptides<sup>55</sup>. Recently, a machine learning approach with increased detection sensitivity, termed XLinkProphet, was developed to model multiple types of information about cross-linked peptide pairs to compute discriminating probabilities and accurately estimate FDR<sup>56</sup>. This protocol describes the use of Comet (Steps 108–115) and XLinkProphet (Steps 116–118) for identification of cross-linked peptide pairs from large-scale XL-MS data sets. The general data analysis workflow is illustrated in Fig. 3.

### Expertise and equipment

The described protocol uses a diverse array of laboratory techniques, equipment, and data processing strategies. As such it requires access to modern LC–MS instrumentation capable of high-resolution mass analysis (i.e., Q-Exactive Plus from Thermo Scientific), a peptide synthesizer (i.e., Liberty Lite from CEM), and cell culture facilities. It is helpful for researchers following this protocol to have a level of expertise in a number of areas to successfully carry out the protocol. Beyond general wet lab experience, knowledge of cell culture techniques, peptide synthesis, protein sample handling, operation of LC–MS instrumentation, and proteomics data analysis will help to ensure success of



**Fig. 3 | Data analysis workflow.** **a**, Mass spectra containing the mass of the intact cross-linked peptide pair, masses of the individual released peptides, and fragment ions of the released peptides are acquired by ReACT or mango methods. **b**, Database search is carried out by Comet to assign peptide sequences to the mass spectra. **c**, PeptideProphet is run on the Comet search results to assign probabilities to individual peptide sequences. **d**, XLinkProphet is run to validate the cross-linked peptide pair and corresponding proteins by assigning a cross-link probability and estimation of the FDR. **e**, Validated results are uploaded into the XLinkDB database, where the identified cross-linked peptide pairs can be viewed in a table or interaction-network format. XLinkDB automatically maps cross-links to existing PDB structures and homology-based structural models and generates distance information for each cross-link. The resulting protein structures with cross-links mapped onto them can be viewed within XLinkDB.

experiments using this protocol. That said, in our laboratory we have successfully used this protocol to train undergraduate and collaborating researchers with limited previous experience in these areas to carry out large-scale cross-linking.

### Limitations

The application of XL-MS to complex biological samples provides low-resolution structural information on proteins and protein complexes. This information can be used to guide molecular modeling and docking algorithms and can be combined with structural information from complementary structural biology techniques such as X-ray crystallography, NMR, and cryo-EM to produce models of protein structures. The use of cross-linker molecules with two reactive groups produces binary interaction information between two reactive amino acid residues. Therefore, additional information from other sources, such as the canonical subunit composition of protein complexes, is needed to extrapolate the XL-MS information to higher-order structural information. Furthermore, cross-linker molecules can react with an ensemble of proteins and complexes present in vivo. In most cases, one cannot be certain that all identified cross-linked peptide pairs on a particular protein were formed at the same time or are compatible with a single structural model or conformation for that protein. Despite great progress being made in recent years, currently, XL-MS studies are still only scratching the surface of the complexity of the interactome, generating on the order of thousands of non-redundant cross-linked peptide pair IDs involving hundreds of proteins. Continued developments will undoubtedly increase the depth and coverage of such studies, providing researchers with valuable new knowledge to address increasingly challenging questions and to improve understanding of biological systems.

## Materials

## Biological materials

▲ **CRITICAL** This protocol uses cells or a tissue sample of interest. The examples listed here have been tested in our lab.

- *Acinetobacter baumannii* Bouvet and Grimont (strain 2208; ATCC, cat. no. 19606). For example use case, see refs. <sup>14,19</sup>.
- *Escherichia coli* (Migula) Castellani and Chalmers (strain K-12; ATCC, cat. no. 10798). For example use case, see refs. <sup>16,36</sup>.
- HeLa cells (ATCC, cat. no. CCL-2). For example use case, see refs. <sup>17,41,42</sup>. **! CAUTION** The cell lines used in your research should be regularly checked to ensure they are authentic and are not infected with mycoplasma.
- H292 cells (ATCC, cat. no. CRL-1848). For example use case, see ref. <sup>14</sup>. **! CAUTION** The cell lines used in your research should be regularly checked to ensure they are authentic and are not infected with mycoplasma.
- Mouse hearts dissected from strain C57BL/6NCrl, male and female, 12–14-week-old animals (Charles River Laboratories, cat. no. IMSR CRL:27). For example use case, see ref. <sup>20</sup>. **! CAUTION** Any experiments involving live animals must conform to relevant institutional and national regulations. All procedures involving animal use were performed with the approval of the Institutional Animal Care and Use Committee of the University of Washington.
- *Pseudomonas aeruginosa* (strain PAO1; University of Washington Transposon Mutant Collection, cat. no. MPAO1). For example use case, see ref. <sup>15</sup>.

## Reagents

**! CAUTION** This protocol uses harmful reagents, including corrosive, toxic and flammable materials. Use appropriate personal protective equipment (PPE), including laboratory coat, gloves, and safety goggles. It is recommended that any transfer of corrosive, toxic, or flammable material be performed in a fume hood. ▲ **CRITICAL** Make sure all solvents and reagents used for synthesis are freshly made and of high purity. Solvents should be ACS grade or higher.

- Acetonitrile (LC–MS grade; Fisher Scientific, cat. no. A955)
- Acetonitrile with 0.1% (vol/vol) TFA acid (LC–MS grade; Fisher Scientific, cat. no. LS121)
- Ammonium bicarbonate ( $\text{NH}_4\text{HCO}_3$ ; Sigma-Aldrich, cat. no. A6141)
- Ammonium hydroxide ( $\text{NH}_4\text{OH}$ ) solution (ACS reagent, 28.0–30.0%  $\text{NH}_3$  basis; Sigma-Aldrich, cat. no. 221228) **! CAUTION** Ammonium hydroxide is corrosive.
- BDP (AnaSpec, cat. no. 66727-2)
- BSA (Sigma-Aldrich, cat. no. A7638)
- Calcium chloride ( $\text{CaCl}_2$ ; Sigma-Aldrich, cat. no. AC42352)
- Dichloromethane (DCM; 99.9% (wt/wt); Fisher Scientific, cat. no. AC610300010) **! CAUTION** DCM is flammable and toxic.
- Diethyl ether (99.5% (wt/wt); Fisher Scientific, cat. no. AC364330010) **! CAUTION** Diethyl ether is flammable and toxic. ▲ **CRITICAL** Diethyl ether can form organic peroxides upon exposure to air and light. Use within 1 year of opening.
- Diisopropylcarbodiimide (DIC; Sigma-Aldrich, cat. no. D125407) **! CAUTION** DIC is flammable, corrosive, and toxic.
- Dimethylformamide (DMF; certified ACS; Fisher Scientific, cat. no. D119) **! CAUTION** DMF is flammable and toxic.
- Dimethyl sulfoxide (DMSO)
- DMEM (Fisher Scientific, cat. no. 11-965-118)
- Ethyl cyano(hydroxyimino)acetate (Oxyma; Sigma-Aldrich, cat. no. 8510860100)
- EDTA (99% (wt/wt); Fisher Scientific, cat. no. AC118432500)
- FBS (Atlas Biologicals, cat. no. FP-0500-A)
- Fmoc-Asp(OtBu)-OH (Millipore Sigma, cat. no. 8520370005)
- Fmoc-Gly-Wang resin (100–200 mesh; Bachem, cat. no. D-1745)
- Fmoc-L-Lys(biotin)-OH (P3 BioSystems, cat. no. 41084)
- Fmoc-Lys(Fmoc)-OH (Millipore Sigma, cat. no. 8520410025)
- Fmoc-Pro-OH (Millipore Sigma, cat. no. 8520170025)
- Formic acid (FA, 99.5+ % (vol/vol), LC–MS grade; Fisher Scientific, cat. no. A117) **! CAUTION** Formic acid is corrosive.



- HEPES (Sigma-Aldrich, cat. no. H3375-100G)
- Iodoacetamide (IAA; Sigma-Aldrich, cat. no. I1149) **! CAUTION** IAA is toxic.
- Magnesium chloride ( $\text{MgCl}_2$ ; Sigma-Aldrich, cat. no. M8266)
- Methanol (LC-MS grade; Fisher Scientific, cat. no. A456) **! CAUTION** Methanol is flammable.
- Monomeric Avidin UltraLink Resin (Pierce Biotechnology, cat. no. 53146)
- *N*-hydroxyphthalimide (NHP; Sigma-Aldrich, cat. no. H53704)
- Nitrogen (g) (99.998%; Praxair, cat. no. NI 4.8-T)
- Penicillin–streptomycin (10,000 U/mL; Fisher Scientific, cat. no. SV30010)
- Phenytoin–pentobarbital solution (Beuthanasia-D Special; Merck, cat. no. NDC:0061-0473) **! CAUTION** Phenytoin–pentobarbital solution is toxic.
- PBS (Fisher Scientific, cat. no. SH30256LS)
- PBS with calcium and magnesium (Fisher Scientific, cat. no. SH30264FS)
- Phosphoric acid (49–51% (wt/vol), HPLC grade; Sigma-Aldrich, cat. no. 79607) **! CAUTION** Phosphoric acid is corrosive.
- Piperidine (99.5% (vol/vol); Sigma-Aldrich, cat. no. 411027) **! CAUTION** Piperidine is flammable and toxic.
- Potassium chloride (KCl; Sigma-Aldrich, cat. no. P9541)
- Potassium phosphate, monobasic ( $\text{KH}_2\text{PO}_4$ ; Sigma-Aldrich, cat. no. P9791)
- Pyridine (DriSolv, anhydrous septum-sealed bottle; Millipore Sigma, cat. no. MPX20127) **! CAUTION** Pyridine is flammable and toxic.
- ReproSil-Pur, 5 micron, 120 Å (Dr. Maisch, cat. no. r15.8e)
- Sodium chloride (NaCl; Sigma-Aldrich, cat. no. S 9888)
- Sodium hydroxide (NaOH; Sigma-Aldrich, cat. no. 221465)
- Sodium phosphate, dibasic ( $\text{Na}_2\text{HPO}_4$ ; Sigma-Aldrich, cat. no. AC20651)
- Sodium phosphate, monobasic ( $\text{NaH}_2\text{PO}_4$ ; Sigma-Aldrich, cat. no. S9638)
- Succinic anhydride (Sigma-Aldrich, cat. no. 239690) **! CAUTION** Succinic anhydride is corrosive and toxic.
- Sucrose
- Trifluoroacetic acid (TFA acid, Optima LC-MS grade; Fisher Scientific, cat. no. A116) **! CAUTION** Trifluoroacetic acid is corrosive.
- Trifluoroacetic anhydride (TFA anhydride; Sigma-Aldrich, cat. no. 106232) **! CAUTION** TFA anhydride is corrosive and toxic. **▲ CRITICAL** TFA anhydride is prone to hydrolysis and should be used within a week of opening.
- Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl; Fisher Scientific, cat. no. 20491) **! CAUTION** TCEP-HCl is corrosive.
- Trypsin (sequencing grade, modified, frozen; Promega, cat. no. V5113)
- Tryptone (pancreatic digest of casein, Bacto; Fisher Scientific, cat. no. DF0123-17-3)
- Urea (Sigma-Aldrich, cat. no. U5378)
- Water (LC-MS grade; Fisher Scientific, cat. no. W6)
- Water with 0.1% (vol/vol) TFA acid (LC-MS grade; Fisher Scientific, cat. no. LS119) **! CAUTION** Trifluoroacetic acid is corrosive
- Yeast extract (dehydrated culture media additive, BD Bacto; Fisher Scientific, cat. no. DF0127-17-9)

## Equipment

- Cryomill (Retsch, cat. no. 20.745.0001)
- Extraction manifold (20 position, 16 × 75-mm tubes; Waters, cat. no. WAT200608)
- Poly-Prep chromatography columns (Bio-Rad, cat. no. 7311550)
- UPLC systems (nanoAcquity; Waters, cat. no. 176816000) or equivalent nano-LC system capable of stable flow in the submicroliter/min range (e.g., Ultimate 3000; Thermo Scientific, cat. no. ULTIM3000RSLCNANO or EASY nLC; Thermo Scientific, cat. no. LC140)
- High resolution mass spectrometer (Orbitrap or Fourier transform ion cyclotron resonance (FTICR) instrument; e.g., Q Exactive Plus; Thermo Scientific, cat. no. IQLAAEGAAPFALGMBDK)
- LC autosampler vials (Fisher Scientific, cat. no. 03-377-299)
- LC autosampler vial caps (Fisher Scientific, cat. no. 13-622-289)
- pH paper (Fisher Scientific, cat. no. 13-640-5070)
- Sep-Pak Vac C18 cartridge, 3cc/500mg (Waters, cat. no. 186004619)
- Vacuum centrifuge (SP Scientific, model no. EZ-2)
- Ultrasonic processor (Cole-Palmer, cat. no. EW-04714-50)

- Thermal mixer (Eppendorf ThermoMixer compact; Sigma-Aldrich, cat. no. T1317)
- Microcentrifuge (Eppendorf, cat. no. 22620444)
- Analytical/preparative HPLC system (Agilent Technologies, cat. no. G1311C)
- SCX column (Luna column; Phenomenex, cat. no. 00G-4398-N0)
- Microcentrifuge tubes (1.5 mL; Fisher Scientific, cat. no. 02-681-320)
- Conical centrifuge tubes (15 and 50 mL; Fisher Scientific, cat. nos. 14-959-53A and 12-565-270)
- Shaker with 15- and 1.5-mL sample blocks (ThermoMixer R; Eppendorf, cat. no. 5355)
- Glass vacuum filter flask (250 mL; Fisher Scientific, cat. no. FB300250)
- Rubber stopper with hole (Fisher Scientific, cat. no. 14-135B)
- Mixer mill (Retch, model no. MM 400)
- 1.5-mL LC autosampler vial (ThermoFisher, cat. no. MSCERT4000-39TR)
- Fritted trapping column (New Objective, cat. no. IF360-100-50-N-5)

#### Software

▲ **CRITICAL** All required software can be run on a standard personal computer equipped with a Linux operating system and at least 4 GB of RAM.

- ReAdW (<https://sourceforge.net/projects/sashimi/files/ReAdW%20%28Xcalibur%20converter%29/>)
- mango (<https://github.com/jpm369/mango>)
- Comet 2018.01 or later (<http://comet-ms.sourceforge.net/>)
- XLinkProphet (<https://github.com/brucelab/xlinkprophet>)
- Perl v.5.24.0+ (<https://www.perl.org/get.html>)

#### Reagent setup

##### Cross-linker cleavage solution

This solution is 95% (vol/vol) TFA acid, 5% (vol/vol) DCM. Add 0.5 mL of DCM to a 15-mL conical tube and adjust volume to 10 mL with TFA acid. This solution must be prepared fresh and used within 1 d.

##### Aspartic acid solution

This solution is 0.2 M Fmoc-Asp(OtBu)-OH in DMF. Weigh out 1.73 g of aspartate (Fmoc-Asp(OtBu)-OH) and dissolve in 21 mL of DMF. This solution must be prepared fresh and used within 2 weeks.

##### Succinic anhydride solution

This solution is 0.2 M succinic anhydride in DMF. Weigh out 0.48 g of succinic anhydride and dissolve in 23 mL of DMF. This solution must be prepared fresh and used within 2 weeks.

##### Biotin lysine solution

This solution is 0.067 M Fmoc-Lys(biotin)-OH in DMF. Weigh out 1.0 g of biotinylated lysine (Fmoc-Lys(biotin)-OH) and dissolve in 32 mL of DMF. This solution must be prepared fresh and used within 2 weeks.

##### Proline solution

This solution is 0.2 M Fmoc-Pro-OH in DMF. Weigh out 1.42 g of proline (Fmoc-Pro-OH) and dissolve in 21 mL of DMF. This solution must be prepared fresh and used within 2 weeks.

##### Branching lysine solution

This solution is 0.2 M Fmoc-Lys(Fmoc)-OH in DMF. Weigh out 1.3 g of lysine (Fmoc-Lys(Fmoc)-OH) and dissolve in 11 mL of DMF. This solution must be prepared fresh and used within 2 weeks.

##### Deprotection solution

This solution is 20% (vol/vol) piperidine in DMF. Measure out 40 mL of piperidine and adjust to a 200-mL total volume with DMF. This solution can be prepared in advance and stored at room temperature (~20–22 °C) for 1 month.

##### Activator solution

This solution is 0.5 M DIC in DMF. Weigh out 6.31 g and dissolve in 100 mL DMF. This solution can be prepared in advance and stored at room temperature for 2 weeks.

#### Activator base solution

This solution is 1 M Oxyma in DMF. Weigh out 14.21 g of Oxyma and dissolve in 100 mL of DMF. This solution can be prepared in advance and stored at room temperature for 1 week.

#### Na<sub>2</sub>HPO<sub>4</sub>, 170 mM, pH 8.0

Dissolve 2.61 g of NaH<sub>2</sub>PO<sub>4</sub> and 0.16 g of Na<sub>2</sub>HPO<sub>4</sub> in 100 mL of deionized water. This buffer can be prepared in advance and stored at room temperature for 7 d.

#### NH<sub>4</sub>HCO<sub>3</sub>: 100 mM, pH 8.0

Dissolve 7.91 g of NH<sub>4</sub>HCO<sub>3</sub> in 1 L of deionized water. Adjust the pH to 8 by adding NH<sub>4</sub>OH solution. This buffer can be prepared in advance and stored at room temperature for 7 d.

#### 0.5 mM EDTA, pH 8.0

Dissolve 18.61 g of EDTA in 80 mL of deionized water and adjust the pH to 8.0 by adding NaOH. Adjust the volume to 100 mL with deionized water. This solution can be prepared in advance and stored at room temperature for 7 d.

#### NaOH, 1.5 M

Dissolve 0.6 g of NaOH in 10 mL of deionized water. This solution can be prepared in advance and stored at room temperature for 1 month.

#### Lysis buffer

Dissolve 480 mg of urea in 1 mL of 100 mM NH<sub>4</sub>HCO<sub>3</sub>. This buffer should be freshly prepared.

#### LB growth medium

Dissolve 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl in 800 mL of deionized water. Adjust to pH 7.0 with 1.5 M NaOH. Adjust the volume to 1 L with deionized water.

#### Mitochondrial isolation medium (MIM) buffer

This solution is 300 mM sucrose, 10 mM HEPES, 0.2 mM EDTA, pH 7.4. Dissolve 10.269 g of sucrose and 0.2383 g of HEPES in 80 mL of deionized water. Add 40 µL of 0.5 mM EDTA solution. Adjust to pH 7.4 with 1.5 M NaOH. Adjust the volume to 100 mL with deionized water. This solution should be freshly prepared and kept at 4 °C.

#### SCX solvent A

This solution is 7 mM KH<sub>2</sub>PO<sub>4</sub>, pH 2.6, 30% (vol/vol) acetonitrile. Dissolve 0.95 g of KH<sub>2</sub>PO<sub>4</sub> in 1 L of 70% deionized water (vol/vol), 30% (vol/vol) acetonitrile, and adjust the pH to 2.6 by adding phosphoric acid. This solution can be prepared in advance and stored at room temperature for 1 month.

#### SCX solvent B

This solution is 7 mM KH<sub>2</sub>PO<sub>4</sub>, pH 2.6, 350 mM KCl, 30% (vol/vol) acetonitrile. Dissolve 0.95 g of KH<sub>2</sub>PO<sub>4</sub> and 24.85 g of KCl in 1 L of 70% deionized water (vol/vol), 30% (vol/vol) acetonitrile, and adjust the pH to 2.6 by adding phosphoric acid. This solution can be prepared in advance and stored at room temperature for 1 month.

#### LC-MS solvent A

This solution is 0.1% (vol/vol) LC-MS-grade FA in LC-MS-grade water. Add 4 mL of LC-MS-grade FA to 4 L of LC-MS-grade water. This solvent can be prepared in advance and stored at room temperature for 1 month.

#### LC-MS solvent B

This solution is 0.1% (vol/vol) LC-MS-grade FA in LC-MS-grade acetonitrile. Add 4 mL of LC-MS-grade FA to 4 L of LC-MS-grade acetonitrile. This solvent can be prepared in advance and stored at room temperature for 1 month.

**PBS**

This solution is 137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>. Add 8.01 g of NaCl, 0.20 g of KCl, 1.42 g of Na<sub>2</sub>HPO<sub>4</sub>, and 0.24 g of KH<sub>2</sub>PO<sub>4</sub> to 1 L of deionized water. This buffer can be prepared in advance and stored at room temperature for 7 d.

**PBS with 20 mM EDTA**

This solution is 137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, 20 mM EDTA. Add 40 mL of 0.5 M EDTA to 1 L of PBS. This buffer can be prepared in advance and stored at room temperature for 7 d.

**Equipment setup****XlinkProphet**

XlinkProphet requires Perl (v.5.24.0+) to run. Perl is a standard release on Linux systems and should not need to be installed. If Perl is already installed, then XlinkProphet requires no further steps in order to be used.

**Comet and ReAdW**

Binaries are available for Comet and ReAdW from their respective repositories; no installation is required.

**mango**

mango has binaries for both Linux and Windows systems. The Windows binary can be used with no modification; setup instructions for the Linux binary are included repository's readme file.

**Procedure****Synthesis of BDP on resin ● Timing 5 h**

**▲ CRITICAL** Steps 1–39 describe the preparation of the cross-linker, which takes 1 d. The BDP cross-linker is a peptide-based molecule synthesized by SPPS. The following steps describe how to synthesize the cross-linker using a CEM Liberty Lite peptide synthesizer. The complete succinylated form (Fig. 1b) can be purchased through any custom peptide synthesis company. We have successfully used BDP obtained from AnaSpec (cat. no. 66727-2). If purchasing the cross-linker on resin, skip to Step 18 for synthesis of the activated NHP esters.

- 1 Weigh out amount of Fmoc-Gly-Wang resin (100–200 mesh) to obtain 0.5 mmol and transfer it to a reaction vessel.  
**▲ CRITICAL STEP** The amount to weigh out depends on resin loading capacity. Typically, the resin has a loading capacity of ~0.8 mmol/g, but this will vary from lot to lot and is specified by the supplier.
- 2 Weigh out 1.73 g of aspartate (Fmoc-Asp(OtBu)-OH) and dissolve it in 21 mL of DMF.
- 3 Weigh out 23 g of succinic anhydride and dissolve in 23 mL of DMF.
- 4 Weigh out 32 g of biotinylated lysine (Fmoc-Lys(biotin)-OH) and dissolve it in 32 mL of DMF.  
**▲ CRITICAL STEP** Fmoc-Lys(biotin)-OH has limited solubility in DMF. The solution can be divided into three ~10.6-mL aliquots in 15-mL centrifuge tubes and heated to 60 °C with constant mixing at 700 r.p.m. on a ThermoMixer for 5–10 min or until dissolved.
- 5 Weigh out 1.42 g of proline (Fmoc-Pro-OH) and dissolve in 21 mL of DMF.
- 6 Weigh out 1.3 g of lysine (Fmoc-Lys(Fmoc)-OH) and dissolve it in 11 mL of DMF.
- 7 Fill the main wash bottle with at least 400 mL of DMF.
- 8 Fill the deprotection bottle with at least 60 mL of deprotection solution.
- 9 Fill the activator bottle with at least 40 mL of activator solution.
- 10 Fill the activator base bottle with at least 20 mL of 1 M Oxyma in DMF.
- 11 Run the CEM synthesizer with the coupling steps listed in Table 1.
- 12 Transfer the resin-containing BDP from the reaction vessel to a 50-mL Eppendorf tube. Use 1–5 mL of DMF to assist with quantitative transfer of resin.

**Synthesis of N-hydroxyphthalimide trifluoroacetate ester (TFA-NHP) ● Timing 4 h**

- 13 Weigh out 5.86 g of NHP (163.139 g/mol) and place it in a 50-mL round-bottom flask.  
**▲ CRITICAL STEP** N-hydroxysuccinimide can be substituted for NHP here, if desired.
- 14 Add a fourfold molar excess of TFA anhydride (210.03 g/mol, 1.511 g/mL = 20 mL).

**Table 1 | CEM synthesizer coupling steps**

Step	Amino acid	Mass (g)	Volume (mL)	Cycle name	Notes
1	Gly-Resin	0.63		0.5-mmol resin swelling	
2	Fmoc-Lys(biotin)-OH	1	32	0.5-mmol triple coupling	Biotin lysine is coupled using a triple-coupling step due to larger volume (32 mL) needed to dissolve the Fmoc-Lys(biotin)-OH
3	Fmoc-Lys(Fmoc)-OH	1.3	11	0.5-mmol single coupling	
4	Fmoc-Pro-OH	1.42	21	0.5-mmol double coupling	
5	Fmoc-Asp(OtBu)-OH	1.73	21	0.5-mmol double coupling	
6	Succinic anhydride	23	23	0.5-mmol succinyl coupling	Succinyl coupling step does not require the addition of activators DIC or Oxyma

**! CAUTION** TFA anhydride is highly corrosive. Use only in a fume hood while wearing proper PPE, including gloves, goggles, and a lab coat.

**▲ CRITICAL STEP** Both acid anhydrides and the esterified cross-linker product are moisture sensitive and readily hydrolyze upon exposure to water. Store in a dry environment and prepare with fresh, dry reagents.

15 Allow the reaction to proceed under a dry N<sub>2</sub> atmosphere with constant mixing for 1.5 h.

16 Evaporate the excess TFA anhydride under vacuum until a dry, white, crystalline product is obtained in a quantitative (~100%) yield.

#### ? TROUBLESHOOTING

17 Transfer the product to 2-mL Eppendorf tubes and store it at -20 °C or use for esterification of BDP in Step 24.

**■ PAUSE POINT** The product can be stored at -20 °C for up to 1 year. Avoid freeze–thaw cycles, as the TFA–NHP will react with water.

### Esterification and purification of BDP ● Timing 6 h

18 Swell BDP resin with a minimum saturating volume of DMF (~1 mL of DMF for every 1 mL of resin) for 5 min.

19 Weigh out a 12-fold molar excess (1.55 g for 0.5 mmol of BDP) of TFA–NHP.

20 Dissolve the TFA–NHP in 10 mL of dry pyridine.

21 Immediately transfer the TFA–NHP–pyridine solution to the BDP resin and mix for 20 min at 1,400 r.p.m. at room temperature, using a ThermoMixer.

22 Transfer the reaction mixture to a Poly-Prep chromatography column. The resin can be divided among multiple columns to avoid exceeding 2 mL of resin per column. Typically, three columns are needed for a 0.5-mmol scale batch. Steps 23–31 will take place in the Poly-Prep column.

23 Connect the Poly-Prep column to a vacuum flask, using an appropriate-sized rubber stopper with a hole, and filter away and discard the pyridine solution.

24 Using vacuum filtration, wash the resin three times with 20 mL of DMF. Discard the DMF wash.

25 Remove the column from the vacuum, plug the outlet, and incubate the resin in 10 mL of DMF for 20 min.

26 Using vacuum filtration, wash the resin three times with 20 mL of DCM. Discard the DCM wash.

27 Incubate the resin in 20 mL of DCM for 20 min.

28 Cleave the PIR cross-linker from the resin by incubating it at room temperature for 3 h in 5 mL of cross-linker cleavage solution.

**▲ CRITICAL STEP** Make sure to seal the top and bottom of the Poly-Prep column using the included caps.

29 Remove the caps from the Poly-Prep column, and collect and save the cross-linker cleavage solution in a 15-mL Falcon tube.

**▲ CRITICAL STEP** Some slight pressure may build during the cleavage reaction. Remove the top cap first to relieve excess pressure before removing the bottom cap.

30 Add 5 mL of cross-linker cleavage solution to the resin and incubate for 5 min to wash.

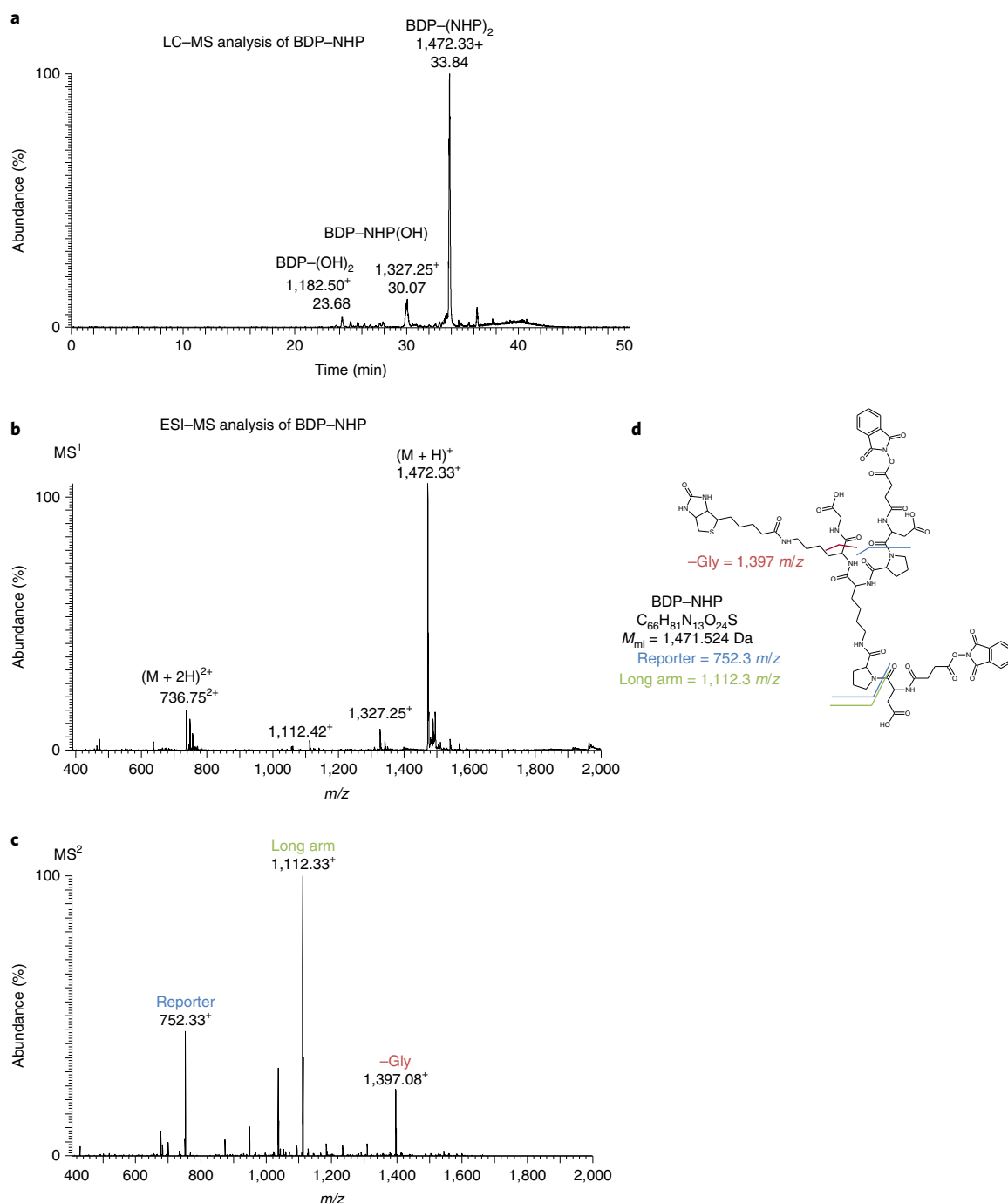
**▲ CRITICAL STEP** Make sure to seal the top and bottom of the Poly-Prep column, using the included caps.



- 31 Remove the caps from the Poly-Prep column and collect the cross-linker cleavage solution and pool it with the solution collected in Step 35.
- 32 Divide 150 mL of ice-cold diethyl ether equally (37.5-mL aliquots) into four 50-mL centrifuge tubes. Slowly add 2.5 mL of cross-linker cleavage solution to each of the four tubes containing the diethyl ether to precipitate the cross-linker.
- 33 Pellet the cross-linker by centrifugation at 3,400g for 30 min at 4 °C.
- 34 Using 5–10 mL of cold diethyl ether per pellet, resuspend the four cross-linker pellets and combine in a single 50-mL tube.
- ▲ **CRITICAL STEP** Use an additional 1–2 mL of ether to wash each tube and ensure quantitative transfer of cross-linker.
- 35 Pellet the cross-linker by centrifugation at 3,400g for 30 min at 4 °C.
- 36 Decant the diethyl ether and dry the cross-linker under vacuum for ~20 min or until completely dry, using a vacuum centrifuge.
- 37 Weigh the product (a >80% yield is expected on the basis of resin loading capacity) and dissolve it in dry DMSO to obtain the desired stock concentration, typically, 200–300 mM.
- ▲ **CRITICAL STEP** Calculate the yield on the basis of the measured mass of product. Convert the measured mass of dried cross-linker into moles by dividing by the average molecular mass of BDP-NHP (1,472.48 g/mol), and then divide by the theoretical maximum number of moles (i.e., 0.5 mmol) on the basis of the amount of resin used. The purity of the cross-linker can be checked by direct infusion ESI-MS analysis of a 1:1,000 dilution of the concentrated stock solution in a solution of 49% (vol/vol) methanol, 49% (vol/vol) water, 2% (vol/vol) acetic acid. A purity of 80–90% should be observed by the relative intensity of the pseudo-molecular ion peak for the PIR cross-linker ( $[M+H]^+ = 1,472.53\ m/z$ ) (Fig. 4).
- ? **TROUBLESHOOTING**
- 38 Measure the concentration of BDP-NHP by UV-Vis absorbance.
- ▲ **CRITICAL STEP** The concentration of BDP-NHP can be measured by UV-Vis absorbance. Dilute a small aliquot (~1 µL) of the concentrated stock solution of cross-linker in 0.1 M  $\text{NH}_4\text{HCO}_3$ , pH 8.0, to achieve an estimated concentration of ~0.5 mM. This will quench the cross-linker and will release free NHP, which has a yellow color. Measure the absorbance of NHP at 300 and/or 410 nm and compare to a calibration curve for free NHP in 0.1 M  $\text{NH}_4\text{HCO}_3$ , pH 8.0, spanning the range from 0.1 to 2 mM to calculate the concentration of released NHP. The concentration of BDP cross-linker in the stock solution is half the calculated concentration of NHP multiplied by the dilution factor.
- 39 Use the cross-linker stock immediately for cross-linking biological samples or divide into aliquots and store at -80 °C for up to 1 year.
- ? **TROUBLESHOOTING**
- **PAUSE POINT** The esterified cross-linker appears to be stable for long period of time when stored at -80 °C; however, repeated freeze-thaw cycles should be avoided because of the absorption of water into the solution, which can hydrolyze the activated esters in the cross-linker.

### Mammalian cell growth ● Timing 5 d

- 40 Establish a culture of cells on appropriate growth medium (DMEM with 10% FBS (vol/vol) and 1% (vol/vol) penicillin-streptomycin. Grow cells in the cell culture incubator operated at 37 °C in a humidified atmosphere containing 5%  $\text{CO}_2$ .
- ▲ **CRITICAL STEP** We have successfully used this approach for HeLa and H292 cells. Growth media and conditions may need to be altered for different cell lines.
- 41 Grow cells in 15-cm plates, using 20 mL of growth medium until cells are 80–90% confluent.
- ▲ **CRITICAL STEP** The above step yields  $\sim 2 \times 10^7$  cells per 15-cm plate and ~5 mg of protein for HeLa cells. Adjust the initial amount of starting material depending on the growth density of the selected cell line.
- 42 Wash the cells with 5 mL of PBS.
- 43 Harvest the cells by adding 5 mL of PBS containing 20 mM EDTA to each plate and returning the plates to the incubator for 5–10 min. Collect the cell suspension by pipetting and pellet by centrifugation at 300g for 3 min at 20 °C.
- ▲ **CRITICAL STEP** Detaching adherent cells with an EDTA-containing buffer is preferred to trypsin to avoid cleavage of cell-surface proteins. If growing cells in a suspension, the PBS-EDTA treatment can be skipped.



**Fig. 4 | LC-ESI-MS analysis of the BDP-NHP cross-linker.** **a**, LC-MS chromatogram of BDP-NHP. The desired product containing two NHP-ester reactive groups elutes at a retention time of 33.84 min with a  $m/z$  value of 1,472.33<sup>+</sup>. The peak at 30.07 min results from hydrolysis of a single NHP group with a  $m/z$  value of 1,327.25<sup>+</sup>. The peak at 23.68 min results from hydrolysis of both NHP groups with a  $m/z$  value of 1,182.50<sup>+</sup>. Some hydrolysis products during this analysis are expected due to the aqueous solvents. **b**, Direct-infusion-generated ESI MS<sup>1</sup> spectrum of BDP-NHP. The single protonated pseudo-molecular ion  $[M+H]^+$  is observed at 1,472.33  $m/z$ , and the doubly protonated form  $[M+2H]^{2+}$  is observed at 736.75  $m/z$ . These peaks should be the most abundant and constitute >80% of the total signal, indicating acceptable purity of the cross-linker. The peak at 1,327.25  $m/z$  represents hydrolysis of one of the NHP reactive groups. The peak at 1,112.42  $m/z$  results from in-source fragmentation of one of the Asp-Pro amide bonds in the cross-linker. **c**, MS<sup>2</sup> spectrum of the precursor ion at 1,472.33  $m/z$ . Major fragment ions are observed at 1,112.33  $m/z$  (long arm) and 752.33  $m/z$  (reporter), resulting from cleavage of one or two of the Asp-Pro amide bonds, respectively. The ion observed at 1,397.08 results from loss of the C-terminal Gly. **d**, Chemical structure of BDP-NHP with the bonds that are cleaved to generate the major fragment ions observed in **c** highlighted. The chemical formula and theoretical monoisotopic mass ( $M_{mi}$ ) are indicated.

- 44 Suspend the cell pellet in PBS with calcium and magnesium.
- 45 Pellet at 300g for 3 min at room temperature and discard the supernatant.
- 46 Resuspend the cell pellet in 10 mL of PBS.
- 47 Pellet at 300g for 3 min at room temperature and discard the supernatant.
- 48 Repeat Steps 46 and 47 two more times, and then proceed immediately to Step 57.

#### Bacterial growth ● Timing 1 d

- 49 Grow bacterial cells at 37 °C to the desired density ( $OD_{600} = 0.6\text{--}1.0$ ) in 250–500 mL of LB growth medium. Transfer to several 50-mL conical tubes.  
**▲ CRITICAL STEP** We have successfully used this approach for *E. coli*, *A. baumannii*, and *P. aeruginosa*. Growth conditions may need to be altered for other bacterial cell types.
- 50 Harvest bacteria by centrifuging at 1,500g for 10 min at room temperature. Discard the supernatant.
- 51 Gently resuspend the bacterial pellet in 30 mL of PBS by pipetting. Centrifuge again at 1,500g for 10 min at room temperature and discard the supernatant. Repeat once more or until the supernatant is clear.  
**▲ CRITICAL STEP** Incomplete washing will leave behind residual growth medium, which contains primary amines that will interfere with cross-linking efficiency.
- 52 Pellet at 1,500g for 10 min at room temperature and discard the supernatant. Proceed immediately to Step 57.

#### Mouse tissue ● Timing 1–3 h

- 53 Anesthetize the mouse with phenytoin–pentobarbital solution at a dose >270 mg/kg intraperitoneally.  
**! CAUTION** Any experiments involving live animals must conform to the relevant institutional and national regulations. All procedures involving animal use were performed with the approval of the Institutional Animal Care and Use Committee of the University of Washington.  
**▲ CRITICAL STEP** Confirm deep anesthesia by lack of response to toe pinch and decreased respiratory rate.
- 54 Excise the tissue of interest and place it into 5 mL of ice-cold MIM buffer.  
**▲ CRITICAL STEP** We have successfully used this approach for heart tissue. Other tissues may need additional washing or preparation before cross-linking.
- 55 Wash all traces of blood from the tissue by gently mixing the tissue in MIM buffer with a glass rod or forceps. Transfer the tissue to a new container with fresh MIM buffer and continue washing until no blood remains.
- 56 Mince the tissue (~150 mg) into 1-mm<sup>3</sup> pieces with a razor blade. Proceed immediately to Step 57.

#### Cross-linking reaction ● Timing 1–2 h

- 57 Resuspend the cell pellet (obtained in Step 48 or 52) or minced tissue sample (Step 56) in one pellet volume (estimated by eye) of 170 mM Na<sub>2</sub>HPO<sub>4</sub> to obtain a 50% (vol/vol) suspension.  
**▲ CRITICAL STEP** Gently pipette up and down several times to resuspend. Pipetting quickly can lead to premature cell lysis.
- 58 Add an appropriate volume of PIR cross-linker stock to reach the desired final concentration, typically a 10 mM final concentration for live-cell cross-linking.  
**▲ CRITICAL STEP** The cross-linker may aggregate after initial addition but will homogenize after shaking for a few minutes.  
**? TROUBLESHOOTING**
- 59 Allow the reaction to proceed for 30 min at room temperature in a ThermoMixer while shaking at 600 r.p.m. The solution will turn yellow as the cross-linker reacts.  
**▲ CRITICAL STEP** The yellow color results from the NHP released from the cross-linker and can be used as a visual indicator of reaction progress. The half-life of the reactive ester groups of the cross-linker is ~8 min in aqueous buffer at pH 8 and ~2 min in the presence of 1 mg/mL BSA. After 30 min, all the reactive cross-linker should have been depleted, and there is no need to quench the reaction.  
**? TROUBLESHOOTING**
- 60 Add one pellet volume of 0.1 M NH<sub>4</sub>HCO<sub>3</sub> to quench any remaining active cross-linker.
- 61 Pellet the cells by centrifugation at 300g for 3 min at 20 °C (1,500g for 10 min for bacterial cells or 1,500g for 3 min for tissue) and remove and discard the supernatant.  
**? TROUBLESHOOTING**

- 62 Resuspend the cells in one pellet volume of fresh 170 mM Na<sub>2</sub>HPO<sub>4</sub>.
- 63 Pellet the cells by centrifugation at 300g for 3 min at 20 °C (1,500g for 10 min for bacterial cells or 1,500g for 3 min for tissue) and remove and discard the supernatant.
- 64 Resuspend cells in one pellet volume of PBS.
- 65 Pellet the cells by centrifugation at 300g for 3 min at 20 °C (1,500g for 10 min for bacterial cells or 1,500g for 3 min for tissue) and remove and discard the supernatant.
- 66 Repeat Steps 64 and 65 until supernatant is no longer yellow, typically about three washes. Proceed directly to Step 67.

### Lysis, protein extraction, and digestion ● Timing 16 h

- 67 Suspend the cross-linked cell pellet by pipetting in lysis buffer.  
**▲ CRITICAL STEP** A minimal volume (0.1–1 mL) should be used to resuspend the cells, as the sample will be diluted tenfold before digestion.
- 68 Lyse the cells by sonication, using an ultrasonic processor until the sample is no longer viscous.  
**▲ CRITICAL STEP** The lysis conditions will depend on the biological sample used. Typically, using a Cole-Palmer ultrasonic processor, five pulses at amplitude 40 for 5 s each is enough to reduce the sample viscosity of cultured human cells. The sample container should be kept in ice to prevent excessive heat buildup. For samples that are difficult to lyse (e.g., tissue samples and Gram-negative bacteria), we use cryogrinding with a cryomill using 5 × 3 min grinding cycles at 30 Hz, to generate a frozen powdered lysate before ultrasonic processing.  
**? TROUBLESHOOTING**
- 69 Reduce protein disulfide bonds by adding TCEP-HCl to a final concentration of 5 mM and incubate at room temperature for 30 min.
- 70 Alkylate reduced thiols with a 10 mM final concentration of IAA for 30 min at room temperature.  
**▲ CRITICAL STEP** IAA is light sensitive, so the container should be shielded from light during the alkylation reaction.
- 71 Dilute the sample tenfold with 0.1 M NH<sub>4</sub>HCO<sub>3</sub> to reduce the urea concentration to <1 M. It may be necessary to transfer the sample to a 15-mL tube.  
**▲ CRITICAL STEP** If the urea concentration is not sufficiently reduced, trypsin digestion will be highly inefficient.
- 72 Perform a Bradford or BCA assay to determine the protein content of the sample. Typical protein concentrations are between 0.5 and 1 mg/mL.
- 73 Add 5 µg of sequencing-grade modified trypsin for every 1 mg of protein in the sample and digest overnight at 37 °C in a ThermoMixer, shaking at 650 r.p.m. Then proceed to Step 74.  
**? TROUBLESHOOTING**

### Desalting ● Timing 4 h

- 74 Acidify digest to pH ~3 by adding TFA acid to a 1% (vol/vol) final concentration.  
**▲ CRITICAL STEP** If the sample pH is <3, then the sample will not bind to the Sep-Pak cartridge, resulting in sample loss.  
**? TROUBLESHOOTING**
- 75 Centrifuge the sample at 16,000g for 15 min at room temperature.
- 76 Select a Sep-Pak of sufficient capacity for the total protein amount determined in Step 72 (protein assay). 50-mg cartridges (column volume = 1 mL) can be used for up to 5 mg of protein, whereas 500-mg cartridges (column volume = 3 mL) can be used for samples exceeding 5 mg of input protein.
- 77 Place the column into the extraction manifold.
- 78 Condition the Sep-Pak by adding 1 column volume of methanol, and drawing it through the column by vacuum.  
**▲ CRITICAL STEP** A pressure difference of 5–10 p.s.i. is sufficient to draw liquid through the cartridge. Higher pressures cause faster flow rates that may inhibit adsorption to the cartridge. The maximum flowrate should be ~1 drop per s. Always verify that the vacuum manifold being used is rated to the pressure that will be applied and monitor the pressure difference using a pressure gauge. Do not allow the cartridge to ever become completely dry. Leave a minimal volume of solvent above the top of the packing material.
- 79 Condition the column by adding 1 column volume of 0.1% (vol/vol) TFA acid in acetonitrile and pulling it through the cartridge. Repeat this step two additional times.

- 80 Equilibrate the column by adding 1 column volume of 0.1% (vol/vol) TFA acid in water and pulling it through the cartridge. Repeat this step two additional times.
- 81 Add supernatant from Step 75 to the column up to maximum volume allowed by the cartridge size and draw it through. Repeat this step until all protein has been added, up to the maximum allowed by the cartridge.
- ▲ **CRITICAL STEP** Collect and save the flow-through from this step. If the pH of the sample is too high, then this flow-through will contain unbound peptides.
- 82 Repeat Step 80 to wash away salts.
- 83 Place an open microcentrifuge tube under the Sep-Pak cartridge outlet to collect the flow-through for the next step.
- 84 Elute peptides from the Sep-Pak cartridge into the open microcentrifuge tube by adding 1 column volume of 80% (vol/vol) acetonitrile/0.1% (vol/vol) TFA acid and drawing it through.
- 85 Concentrate sample to <100  $\mu$ L by vacuum centrifugation.
- **PAUSE POINT** The sample can be stored at -80 °C for up to 1 year.

### Strong cation exchange fractionation ● Timing 2-h initial setup + 2 h per sample

- 86 Dilute the sample to a 500  $\mu$ L volume with SCX solvent A.
- 87 Centrifuge the sample at 16,000g for 15 min at room temperature.
- 88 Transfer the supernatant to a 1.5-mL LC autosampler vial.
- 89 Inject the sample into an LC system equipped with an SCX column.
- ▲ **CRITICAL STEP** Verify that the injection volume is compatible with the sample loop in use on the HPLC system. In addition, verify that the protein content of the injection does not exceed the loading capacity of the column, which is 10 mg for the column specified here.
- 90 Fractionate the sample, using a flow rate of 1.5 mL/min and a 97.5-min gradient of an increasing percentage of SCX solvent B as follows: 0% B at 0 min, 5% B at 7.5 min, 60% B at 47.5 min, 100% B at 67.5 min, 100% B at 77.5 min, 0% B at 77.51 min, and 0% B at 97.5 min. Fractions are collected every 5 min, starting at 17.5 min. Monitor the absorbance of eluting peptides at wavelengths 214 and 280 nm.
- ? **TROUBLESHOOTING**
- 91 Pool the fractions (six pools total) as follows: 1–5, 6–7, 8, 9, 10, and 11–14. Then dry the fraction pools to a final volume of ~2 mL per pooled sample in a vacuum centrifuge. Adjust the pH to 8.0, checked by pH strip, with 1.5 M NaOH.
- ▲ **CRITICAL STEP** Avidin binding efficiency is greatly reduced if the pH is too low.

### Biotin capture ● Timing 4 h

- 92 Enrich the PIR-labeled peptides by adding 200  $\mu$ L of monomeric avidin UltraLink resin to each pooled fraction. Fractions 1–5 are unlikely to contain a substantial number of cross-linked peptides and are typically not subjected to biotin capture.
- 93 Incubate the fractions for 30 min at room temperature on a ThermoMixer (1,000 r.p.m.).
- 94 Place a Poly-Prep column into the extraction manifold for each pooled fraction to be enriched.
- 95 Transfer avidin bound, cross-linked peptides to each empty Poly-Prep column.
- 96 Add 2 mL of 0.1 M  $\text{NH}_4\text{HCO}_3$ , pH 8.0, to the avidin bound, cross-linked peptides in each column, and then remove the liquid by applying vacuum. Repeat this wash five times. The washes can be discarded.
- ▲ **CRITICAL STEP** A pressure difference of 5 p.s.i. is enough to pull liquid through the column.
- 97 Incubate the avidin resin with 1 mL of 70% (vol/vol) acetonitrile, 1% (vol/vol) FA for 5 min.
- 98 Place an empty microcentrifuge tube under each cartridge and collect the flow-through by applying vacuum.
- 99 Concentrate each of the pooled samples (6–7, 8, 9, 10, and 11–14) to ~10  $\mu$ L in a vacuum centrifuge.
- **PAUSE POINT** The samples can be stored at -80 °C for several months.

### LC-MS analysis ● Timing 3 or 5 h per fraction injection, depending on gradient

- ▲ **CRITICAL** Steps 100–107 describe the preparation and LC-MS analysis steps for the enriched cross-linked peptide pair samples. These steps need to be completed for each of the pooled fraction samples (five samples total; fractions 6–7, 8, 9, 10, and 11–14).
- 100 Add 20–50  $\mu$ L of 0.1% (vol/vol) FA to reconstitute cross-linked peptide samples.
- ▲ **CRITICAL STEP** If the peptides were dried to completion, first add a minimal volume of 50% (vol/vol) acetonitrile/50% water/0.1% (vol/vol) FA to re-dissolve the peptides, and then dilute with



- 0.1% (vol/vol) FA to the final desired volume. Ensure that the final acetonitrile concentration is <5% (vol/vol) to ensure efficient trapping and separation by reverse-phase LC.
- 101 Centrifuge at 16,000g for 10 min at room temperature to pellet any particulate.
  - 102 Transfer the supernatant to a LC-MS vial.
  - 103 Inject 1–5 µL of sample into the nano-LC system.
  - 104 Load the peptides onto a trap column (3 cm × 100-µm i.d., fused-silica trap column packed with a stationary phase consisting of ReproSil-Pur C8 (5-µm diameter and 120-Å-pore-size particles)) with a flow rate of 2 µL/min of mobile phase consisting of 98% (vol/vol) LC-MS solvent A (H<sub>2</sub>O containing 0.1% (vol/vol) FA) and 2% (vol/vol) LC-MS solvent B (acetonitrile containing 0.1% (vol/vol) FA) for 10 min.
  - 105 Fractionate the peptides over a 60 cm × 75-µm i.d., fused-silica analytical column packed with ReproSil-Pur C8 (5-µm diameter and 120-Å-pore-size particles) by applying a linear gradient from 95% LC-MS solvent A, 5% LC-MS solvent B to 60% LC-MS solvent A, 40% LC-MS solvent B over either 120 or 240 min at a flow rate of 300 nL/min. The analytical column can be heated to 40 °C to enhance the separation if a column heater is available.
  - 106 Ionize the eluting peptides by ESI with a voltage of 2–2.5 kV to the laser-pulled fused-silica spray tip at the end of the column.
- ? TROUBLESHOOTING**
- 107 Operate the mass spectrometer using a data-dependent analysis method in which one high-resolution MS1 (70,000 resolving power at 200 *m/z*; automated gain control (AGC) = 1 × 10<sup>6</sup>) is followed by 20 high-resolution MS2 scans selecting ions of 4+ to 8+ (70,000 resolving power at 200 *m/z*, AGC = 5 × 10<sup>4</sup>, maximum ion time = 100 ms, isolation window = 3 Da, collision energy = 30 normalized collision energy (NCE), dynamic exclusion = 30 s).
- ? TROUBLESHOOTING**

#### Cross-link search ● Timing 5–60 min per fraction

- 108 Convert the instrument data files (e.g., .raw files) from Step 107 to the open .mzXML format. This can be done using either vendor-specific software or a tool such as msconvert or ReAdW, available as part of the Trans Proteomic Pipeline<sup>57</sup>.
- 109 Download the current release of mango from <https://github.com/jpm369/mango>. The repository includes 64-bit binaries for both Windows and Linux operating systems, as well as detailed usage instructions and parameter descriptions.
- 110 Generate a mango.params.new file by typing the following command into a Terminal window:  
`mango.exe --p`
- 111 Adjust the settings in the mango.params.new file using a text editor such as nano or notepad, save it, and then rename it as mango.params. The default parameters in the file work for BDP-NHP cross-linked samples run using the mass spectrometer method described in Step 107.
- 112 Run mango.exe on a .mzXML file; this will return a .ms2 file containing individual precursor masses of released peptides for each spectrum, as well as a .peaks file containing all relationships within tolerance. Use the following command:  
`mango.exe file.mzXML`
- 113 Using the included Comet binary in the mango GitHub repository, generate a comet.params.new file. Use the following command:  
`comet.exe -p`
- 114 Use a text editor to adjust the settings in the comet.params.new file to match your experiment. A sample high-resolution comet.params file is included in the TestCase folder of the mango repository. Details on the various parameters used by Comet can be found at [http://comet-ms.sourceforge.net/parameters/parameters\\_201801/](http://comet-ms.sourceforge.net/parameters/parameters_201801/). Set the output\_txtfile parameter to 1 to return an easy-to-parse output. Once the parameters have been set, rename the file to comet.params.
- 115 Run comet.exe on the .ms2 file generated by mango (from Step 112); this will return a .pep.xml file containing unpaired results. Use the following command:  
`comet.exe file.ms2`

#### Post-processing ● Timing 5–60 min per fraction

- 116 Run PeptideProphet<sup>58</sup> on the Comet search results from Step 115 to validate with assigned probabilities. Use the command: `xinteract -OEAdP -p0 -PPM -l5 search_result1.pep.xml search_result2.pep.xml, ..., where ‘...’ indicates listing all .pep.xml file search`

result files to be included in the analysis. Note that when the search database has decoys with a unique protein prefix (e.g., 'rev\_'), that information should be passed as an additional parameter `-dprefix` (e.g., `-drev_`). Use the following command: `xinteract -OEAdP -p0 -PPM -15 -dprefix search_result1.pep.xml search_result2.pep.xml, ...`

- 117 Run iProphet<sup>59</sup> to further refine PeptideProphet probabilities on the basis of additional peptide-specific information, including the number of replicate spectra, number of charge states a peptide was observed at, and different modified forms of a peptide. Use the following command:

```
InterProphetParser interact.pep.xml iprophet.pep.xml
```

- 118 Run XLinkProphet<sup>56</sup> (available at <https://github.com/brucelab/xlinkprophet>) to validate mango results based on iProphet search result probabilities and cross-link properties. Note that the mango peaks files should be in the same directory locations as the Comet search result pepXML files. Use the following command:

```
XLinkProphet.pl iprophet.pep.xml MIN_PEPLEN=5 CROSSLINK_MODS=K:329.152492 REPORTERMASS=xxxx
```

where K:329.152492 indicates the amino acid and modification mass of the cross-linker stump used to define the site of cross-link attachment to the peptide (this should be substituted with a value or comma-delimited values appropriate for the analyzed data), and xxxx is to be substituted by the actual mass of the cross-linker reporter relevant to the data (default = 751.40508 for BDP).

### ? TROUBLESHOOTING

#### Automated interaction network and structural modeling/docking ● Timing 1 h

- 119 Upload the resulting data from XLinkProphet (Step 118) into XLinkDB at <http://xlinkdb.gs.washington.edu/>.
- 120 Arrange the cross-link identifications in a tab-delimited text file with the following six columns: Peptide A | Protein A | Labeled position A || Peptide B | Protein B | Labeled position B. A template file is available at <http://xlinkdb.gs.washington.edu/xlinkdb/Help.php>.
- 121 Upon successful upload, an interactive interaction network will be generated, and cross-linked proteins will be automatically submitted for structural modeling and molecular docking algorithms producing structural models in PDB format, which are viewable directly from the website.

## Troubleshooting

Troubleshooting advice can be found in Table 2.

**Table 2 | Troubleshooting table**

Step	Problem	Possible reason	Solution
16	Low yield (<50%) of TFA-NHP	Degraded or contaminated reagents used during synthesis	Use fresh, high-purity reagents. TFA anhydride is prone to hydrolysis and should be used within a week of opening
37	Low yield (<50%) of cross-linker during resin cleavage or resin has uncleaved cross-linker remaining	Residual DMF on cross-linker resin inhibits cleavage	Wash the resin extensively with DCM, and then incubate the resin in DCM for 20 min, pipetting occasionally to improve solvent accessibility. Make a fresh stock of cross-linker cleavage solution and repeat the incubation in cleavage solution
	Impure cross-linker observed during direct infusion ESI-MS analysis	Degraded or contaminated reagents used during synthesis	Make sure all solvents and reagents used for synthesis are freshly made and of high purity. Solvents should be of ACS grade or higher. DMF should not be >6 months old. Amino acid and DIC solutions are stable for up to 2 weeks. The piperidine solution is stable for 1 month
39	Cross-linker has hydrolyzed	Cross-linker stock solution has become contaminated with water	Verify hydrolysis by ESI-MS analysis (Fig. 4)
58	Cross-linker precipitated in pipette tip	Pipette tip was immersed in aqueous buffer before expelling the concentrated cross-linker solution, causing it to precipitate	Do not immerse the pipette tip in the solution when adding cross-linker. Instead, hold the tip above the solution and proceed with pushing the cross-linker solution out

Table continued

Table 2 (continued)

Step	Problem	Possible reason	Solution
59	No yellow color observed during cross-linking reaction	pH of solution is too low	Check the pH of the reaction solution. It should be between 7 and 8. If it is lower, adjust it to within range by titrating with 1 M NaOH or use a higher concentration of buffer
61	Cell pellet is small relative to pellet size before cross-linking (~50% or more reduction in size) or solution is sticky and viscous	Cells have lysed during pelleting and resuspension	Check centrifugation speed to make sure it is appropriate for the sample type used. Be extremely gentle while using a pipette to resuspend cells, only using a minimal number of cycles to suspend the cell pellet
68	Tissue samples not lysing during cryogrinding	Cryogrinding ball was added to liquid sample before freezing in liquid nitrogen and is now encased in ice and not able to grind sample	Add cryogrinding ball after sample solution has been frozen in liquid N <sub>2</sub> to ensure it is free to move during shaking
73	Large amount of precipitation in tryptic-digest sample	Trypsin failed to digest proteins into peptides	Sample solution conditions were not compatible with trypsin. Check the pH to ensure it is between 7 and 8. Ensure urea concentration is <1 M. Correct solution conditions and add fresh trypsin to repeat Step 71
74	Large amount of precipitation observed after acidifying sample	Large amount of residual undigested protein due to inefficient trypsin digestion	Check solution conditions for trypsin digestion. Make sure the pH is between 7 and 8, urea concentration is <1 M and fresh trypsin is used. Repeat Steps 71 and 72
90	No signal or low signal (<100 milliabsorbance units at 214 nm) observed in SCX absorbance traces	Loss of protein during desalting due to poor binding to the packing material	Check the pH of the input sample for the Sep-Pak. Re-acidify if >3, and repeat the desalting procedure
106	Spray is sputtering or droplets form at the tip of the column	Incorrect spray voltage or tip position; blunt, broken, or dirty spray tip	Adjust the applied spray voltage and tip position until stable spray is obtained. If the tip appears blunt or dirty, replace with a new column
107	No signal or low signal (<10 <sup>6</sup> absolute intensity in the total ion chromatogram on a Q-Exactive Plus mass spectrometer) observed in chromatogram on mass spectrometer	Poor binding of sample to monomeric avidin resin	Verify that the pH of the SCX fractions after base addition are ~8. These samples can be recaptured again after the pH is corrected
		Sample not injected by autosampler on LC	Verify that autosampler is delivering the requested volume of sample onto the LC column
		Mass spectrometer is out of calibration or has low sensitivity due to dirty ion optics	Clean and calibrate the mass spectrometer, verifying sensitivity with a standard sample
118	No correct identifications in .pep.xml file	Multiple contributing factors	Use a cross-linked purified protein such as BSA as a positive-control sample and ensure that results similar to those presented in Fig. 5 and Supplementary Table 1 are obtained
		Incorrect database or search parameters	Check that the correct protein database was used and the Comet search parameters match the experimental setup

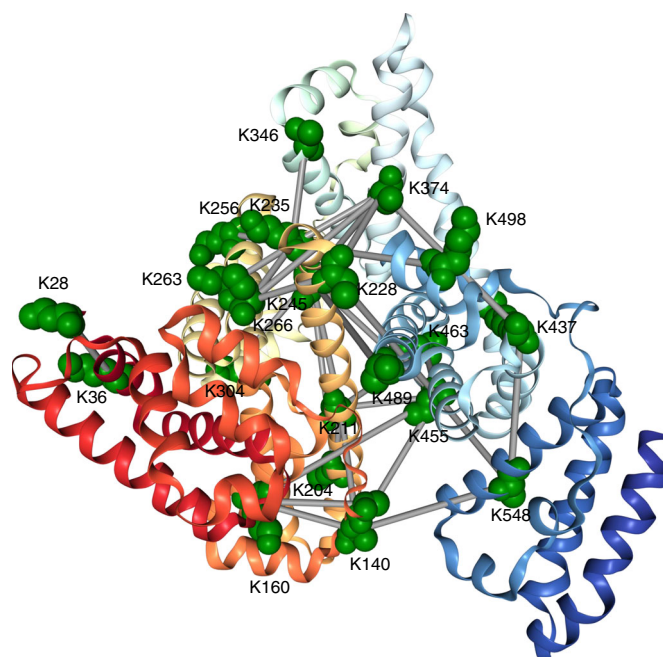
## Timing

Steps 1–12, synthesis of BDP on resin: 5 h  
 Steps 13–17, synthesis of TFA-NHP: 4 h  
 Steps 18–39, esterification and purification of BDP: 6 h  
 Steps 40–48, mammalian cell growth: 5 d  
 Steps 49–52, bacterial growth: 1 d  
 Steps 53–56, mouse tissue: 1–3 h  
 Steps 57–66, cross-linking reaction: 1–2 h  
 Steps 67–73, lysis, protein extraction, and digestion: 16 h  
 Steps 74–85, desalting: 4 h  
 Steps 86–91, strong cation exchange fractionation: 2-h initial setup + 2 h per sample  
 Steps 92–99, biotin capture: 4 h  
 Steps 100–107, LC–MS analysis: 3 or 5 h per fraction injection, depending on gradient  
 Steps 108–115, cross-link search: 5–60 min per fraction  
 Steps 116–118, post-processing: 5–60 min per fraction  
 Steps 119–121, automated interaction network and structural modeling/docking: 1 h

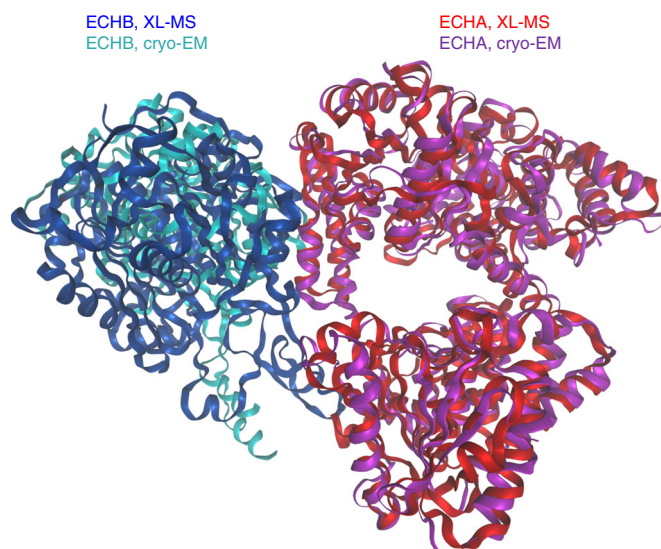
## Anticipated results

The workflow described in this protocol can be used to synthesize a PIR cross-linker and make large-scale structural biology measurements of protein structures and interactions within cells. The cross-linker should be obtained in high purity (>80%) as evaluated by LC-ESI-MS and direct infusion ESI-MS analysis (Fig. 4). Successful use of this protocol will result in the identification of thousands of non-redundant cross-linked peptide pairs resulting from hundreds of proteins when filtered to an estimated FDR of  $\leq 1\%$ . We routinely observe that between 70 and 80% of the identified *in vivo* cross-linked peptide pairs are intra-protein links, whereas 20–30% are inter-protein links. This observation can be explained by the fact that <100% of the total cellular abundance of a given protein will exist in a complex with another protein in the cell. Therefore, for a given cross-linked Lys, proximal intra-protein Lys residues will be present at a much higher frequency than that of inter-protein residues. Despite the observation that many non-redundant cross-linked peptide pairs are identified from highly abundant proteins, cross-linked peptides from proteins localized in all cellular compartments are normally identified. Moreover, cross-linked peptides from lower-abundance proteins are often identified, albeit with fewer numbers of pairs. For example, when comparing 3,323 cross-linked peptide pairs identified from 893 proteins in HeLa cells against protein copy-number data<sup>60</sup>, we found that 90% of the cross-linked proteins were in the top 3,263 most abundant proteins. However, the cross-linked proteins spanned a wide range in abundance, from  $10^8$  copies per cell down to  $10^2$  copies per cell, with a median of  $7.8 \times 10^5$  copies per cell<sup>41</sup>. As for experimental controls, a non-cross-linked control sample should not yield any correct cross-linked peptide pair identifications. There should be minimal signal observed in the LC-MS analysis resulting from non-cross-linked peptides that non-specifically bind to the monomeric avidin resin. XLinkProphet will send an e-mail summarizing the results, which in the case of a negative-control sample should read: 'Warning: No correct crosslinks found. Please check your run parameters and number of correct search results in your input files.' Results from cross-linked BSA should yield results similar to the 37 non-redundant cross-linked peptide pairs listed in Supplementary Table 1 and displayed on the structure of BSA in Fig. 5. If the anticipated results are not obtained, consult the Troubleshooting table (Table 2).

Comparison of the identified cross-linked sites with existing structural data in the PDB results in high agreement, with ~95% of the links having Euclidean Ca–Ca distances of <35 Å (ref. <sup>41</sup>). This observation indicates that cross-linking-derived distance constraints can be useful for generating structural models for proteins and protein complexes for which there is little to no existing structural



**Fig. 5 | Cross-links mapped onto BSA structure.** The cross-links listed in Supplementary Table 1 are mapped on the structure (PDB 3V03). The backbone is displayed as a ribbon colored with a gradient from red at the N terminus to blue at the C terminus. Cross-linked lysine residues are displayed as green space-filled residues and cross-links between them are shown as gray bars.



**Fig. 6 | Comparison of structural models for the mitochondrial trifunctional enzyme complex generated by XL-MS guided docking and cryo-EM.** The trifunctional enzyme consists of alpha (ECHA, red = XL-MS; magenta = cryo-EM) and beta (ECHB, dark blue = XL-MS; teal blue = cryo-EM) subunits. The XL-MS interaction model was generated by cross-link-guided molecular docking using 28 Lys-Lys distance restraints identified from large-scale XL-MS analysis of mouse heart tissue<sup>20</sup>; the cryo-EM model is the human sequence. PDB 5ZQZ was published by Liang et al.<sup>61</sup>.

data. Cross-link-guided molecular docking utilizing distance restraint data from large-scale XL-MS experiments can produce molecular models for protein complexes that closely resemble models produced by other structural biology techniques, including cryo-EM. For example, large-scale cross-linking data resulting from XL-MS analysis of intact mitochondria produced a cross-link-derived model for the mitochondrial oxidative phosphorylation (OXPHOS) supercomplex consisting of complex I and complex III. This model was found to be within a 1.3- to 2.4-Å root mean square deviation (r.m.s.d.) of a cryo-EM-derived orthologous model<sup>10</sup>. In another example, 28 cross-linked peptide pairs identified by XL-MS analysis of mouse heart tissue were used to construct a molecular model for the interaction between the mitochondrial trifunctional enzyme alpha subunit (ECHA) and beta subunit (ECHB), each of which is important for beta oxidation of fatty acids<sup>20</sup>. The cross-link-directed ECHA-ECHB model structure was observed to be very similar (r.m.s.d. = 3.6 Å for ECHA; 7.9 Å for ECHB) to subsequent model structure—resultant cryo-EM measurements of the human ortholog complex<sup>61</sup> (Fig. 6).

In vivo XL-MS can also provide structural and interaction information on previously unknown complexes. In bacterial systems, in vivo XL-MS measurements enabled identification of protein interactions important in both virulence and drug resistance in a clinical isolate of ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *A. baumannii*, *P. aeruginosa*, and *Enterobacter* spp.), which are the leading cause of hospital-acquired infections throughout the world. For instance, in vivo XL-MS studies with *A. baumannii* AB-5075, isolated from a patient at Walter Reed Army Medical Center<sup>62</sup>, identified interactions between the carbapenemase Oxa23 and a series of outer-membrane porin proteins, including carbapenem-associated resistance protein (CarO). These cross-linked peptides uncovered new mechanistic details on increased carbapenem resistance in this multidrug-resistant, highly virulent pathogen<sup>19</sup>. XL-MS can also be used to characterize host–pathogen molecular interactions, as demonstrated in a model of bacterial infection using *A. baumannii* AB-5075 and the human bronchial epithelial cell line H292. In those studies, a key virulence factor in *A. baumannii*, outer-membrane protein A (OmpA) was identified as cross-linked to nuclear, mitochondrial and desmosomal proteins<sup>14</sup>. These are a few specific examples demonstrating the potential to utilize large-scale XL-MS data to derive useful new structural data on protein complexes as they reside within their native biological systems. However, as more labs begin to implement in vivo cross-linking experiments with this or other protocols, the utility of identified cross-linked peptides will grow beyond a single research lab or even beyond a single research field. Once identified, each cross-linked peptide pair can serve as a useful in vivo probe for a particular protein conformation or interaction. This information can be exploited in targeted quantitative experiments using MS methods such as parallel reaction monitoring (PRM)<sup>40</sup> to quantify cross-link levels between differing biological states or perturbations such as drug treatment in cells. The generality of this approach was demonstrated in a



cross-laboratory study utilizing PRM to measure drug-induced structural and interaction changes of heat shock protein 90 (Hsp90)<sup>40</sup>. Thus, each identified cross-linked peptide pair resultant from application of this protocol will not only provide the individual researcher with new insight on his or her own biological questions and further increase the collective knowledge base of in vivo interactomes, but also contribute to a massive array of PRM assays useful for in vivo studies of thousands of proteins and complexes with any type of perturbation.

### Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Data are freely available online. Raw data from Chavez, J. D. et al. *Cell Syst.* **6**, 136–141 (2018) <https://doi.org/10.1016/j.cels.2017.10.017> (used for Fig. 6) are available in the PRIDE Archive (<http://www.ebi.ac.uk/pride/archive/>) with accession no. PXD007673. Cross-linking data are available in XLinkDB (<http://xlinkdb.gs.washington.edu/>) under network name ChavezCellSystems2017 BruceLab.

### Code availability

All necessary software is freely available for the foreseeable future online at the respective sources: XLinkProphet (<https://github.com/brucelab/xlinkprophet>), mango (<https://github.com/jpm369/mango>), Comet (<http://comet-ms.sourceforge.net/>), ReAdW (<https://sourceforge.net/projects/sashimi/files/ReAdW%20%28Xcalibur%20converter%29/>), and perl (<https://www.perl.org/get.html>). In addition, the versions of these software used in this protocol are available in the Supplementary Software file.

## References

1. Holding, A. N. XL-MS: protein cross-linking coupled with mass spectrometry. *Methods* **89**, 54–63 (2015).
2. Leitner, A. et al. Crosslinking and mass spectrometry: an integrated technology to understand the structure and function of molecular machines. *Trends Biochem. Sci.* **41**, 20–32 (2016).
3. Paramelle, D. et al. Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* **13**, 438–456 (2013).
4. Sinz, A. Divide and conquer: cleavable cross-linkers to study protein conformation and protein–protein interactions. *Anal. Bioanal. Chem.* **409**, 33–44 (2017).
5. Rappsilber, J. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540 (2011).
6. Anderson, G. A. et al. Informatics strategies for large-scale novel cross-linking analysis. *J. Proteome Res.* **6**, 3412–3421 (2007).
7. Leitner, A. et al. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell Proteomics* **9**, 1634–1649 (2010).
8. Liu, F. et al. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).
9. Tan, D. et al. Trifunctional cross-linker for mapping protein–protein interaction networks and comparing protein conformational states. *Elife* **5**, e12509 (2016).
10. Schweppe, D. K. et al. Mitochondrial protein interactome elucidated by chemical cross-linking mass spectrometry. *Proc. Natl. Acad. Sci. USA* **114**, 1732–1737 (2017).
11. Liu, F. et al. The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for coexisting respiratory supercomplexes. *Mol. Cell Proteomics* **17**, 216–232 (2018).
12. Fasci, D. et al. Histone interaction landscapes visualized by crosslinking mass spectrometry in intact cell nuclei. *Mol. Cell Proteomics* **17**, 2018–2033 (2018).
13. Wu, X. et al. Dynamic proteome response of *Pseudomonas aeruginosa* to tobramycin antibiotic treatment. *Mol. Cell Proteomics* **14**, 2126–2137 (2015).
14. Schweppe, D. K. et al. Host–microbe protein interactions during bacterial infection. *Chem. Biol.* **22**, 1521–1530 (2015).
15. Navare, A. T. et al. Probing the protein interaction network of *Pseudomonas aeruginosa* cells by chemical cross-linking mass spectrometry. *Structure* **23**, 762–773 (2015).
16. Weisbrod, C. R. et al. In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. *J. Proteome Res.* **12**, 1569–1579 (2013).
17. Chavez, J. D. et al. Protein interactions, post-translational modifications and topologies in human cells. *Mol. Cell Proteomics* **12**, 1451–1467 (2013).
18. Kaake, R. M. et al. A new in vivo cross-linking mass spectrometry platform to define protein–protein interactions in living cells. *Mol. Cell Proteomics* **13**, 3533–3543 (2014).

19. Wu, X. et al. In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in *Acinetobacter baumannii* strain AB5075. *Nat. Commun.* **7**, 13414 (2016).
20. Chavez, J. D. et al. Chemical crosslinking mass spectrometry analysis of protein conformations and super-complexes in heart tissue. *Cell Syst.* **6**, 136–141.e5 (2018).
21. Klykov, O. et al. Efficient and robust proteome-wide approaches for cross-linking mass spectrometry. *Nat. Protoc.* **13**, 2964–2990 (2010).
22. Iacobucci, C. et al. A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein–protein interactions. *Nat. Protoc.* **13**, 2864–2889 (2018).
23. Orban-Nemeth, Z. et al. Structural prediction of protein models using distance restraints derived from cross-linking mass spectrometry data. *Nat. Protoc.* **13**, 478–494 (2018).
24. Tang, X. et al. Mass spectrometry identifiable cross-linking strategy for studying protein–protein interactions. *Anal. Chem.* **77**, 311–318 (2005).
25. Tang, X. & Bruce, J. E. A new cross-linking strategy: protein interaction reporter (PIR) technology for protein–protein interaction studies. *Mol. Biosyst.* **6**, 939–947 (2010).
26. Petrotchenko, E. V. et al. BiPS, a photocleavable, isotopically coded, fluorescent cross-linker for structural proteomics. *Mol. Cell Proteomics* **8**, 273–286 (2009).
27. Yang, L. et al. In vivo application of photocleavable protein interaction reporter technology. *J. Proteome Res.* **11**, 1027–1041 (2012).
28. Zheng, Q. et al. Probing protein 3D structures and conformational changes using electrochemistry-assisted isotope labeling cross-linking mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **27**, 864–875 (2016).
29. Zheng, Q. et al. Cross-linking electrochemical mass spectrometry for probing protein three-dimensional structures. *Anal. Chem.* **86**, 8983–8991 (2014).
30. Kao, A. et al. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell Proteomics* **10**, M110.002212 (2011).
31. Muller, M. Q. et al. A universal matrix-assisted laser desorption/ionization cleavable cross-linker for protein structure analysis. *Rapid Commun. Mass. Spectrom.* **25**, 155–161 (2011).
32. Petrotchenko, E. V., Serpa, J. J. & Borchers, C. H. An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell Proteomics* **10**, M110.001420 (2011).
33. Tang, X. et al. Profiling the membrane proteome of *Shewanella oneidensis* MR-1 with new affinity labeling probes. *J. Proteome Res.* **6**, 724–734 (2007).
34. Schmidt, C. & Urlaub, H. Combining cryo-electron microscopy (cryo-EM) and cross-linking mass spectrometry (CX-MS) for structural elucidation of large protein assemblies. *Curr. Opin. Struct. Biol.* **46**, 157–168 (2017).
35. Rinner, O. et al. Identification of cross-linked peptides from large sequence databases. *Nat. Methods* **5**, 315–318 (2008).
36. Mohr, J. P. et al. Mango: a general tool for collision induced dissociation-cleavable cross-linked peptide identification. *Anal. Chem.* **90**, 6028–6034 (2018).
37. Zheng, C., Perumalla, P., Chavez, J. D., Eng, J. K. & Bruce, J. E. XLink-DB: database and software tools for storing and visualizing protein interaction topology data. *J. Proteome Res.* **12**, 1989–1995 (2013).
38. Schweppe, D. K. et al. XLinkDB 2.0: integrated, large-scale structural analysis of protein crosslinking data. *Bioinformatics* **32**, 2716–2718 (2016).
39. Keller, A., Chavez, J. D., Eng, J. K., Thornton, Z. & Bruce, J. E. Tools for 3D interactome visualization. *J. Proteome Res.* **18**, 753–758 (2018).
40. Chavez, J. D. et al. A general method for targeted quantitative cross-linking mass spectrometry. *PLoS ONE* **11**, e0167547 (2016).
41. Chavez, J. D., Schweppe, D. K., Eng, J. K. & Bruce, J. E. In vivo conformational dynamics of Hsp90 and its interactors. *Cell Chem. Biol.* **23**, 716–726 (2016).
42. Chavez, J. D. et al. Quantitative interactome analysis reveals a chemoresistant edgotype. *Nat. Commun.* **6**, 7928 (2015).
43. Zhong, X. et al. Large-scale and targeted quantitative cross-linking MS using isotope-labeled protein interaction reporter (PIR) cross-linkers. *J. Proteome Res.* **16**, 720–727 (2016).
44. Chavez, J. D. et al. Cross-linking measurements of the potato leafroll virus reveal protein interaction topologies required for virion stability, aphid transmission, and virus–plant interactions. *J. Proteome Res.* **11**, 2968–2981 (2012).
45. DeBlasio, S. L. et al. Visualization of host–poliovirus interaction topologies using protein interaction reporter technology. *J. Virol.* **90**, 1973–1987 (2016).
46. Alexander, M. M. et al. Insights in luteovirid structural biology guided by chemical cross-linking and high resolution mass spectrometry. *Virus Res.* **241**, 42–52 (2017).
47. Zhang, H. et al. In vivo identification of the outer membrane protein OmcA–MtrC interaction network in *Shewanella oneidensis* MR-1 cells using novel hydrophobic chemical cross-linkers. *J. Proteome Res.* **7**, 1712–1720 (2008).
48. Ramsey, J. S. et al. Protein interaction networks at the host-microbe interface in *Diaphorina citri*, the insect vector of the citrus greening pathogen. *R. Soc. Open Sci.* **4**, 160545 (2017).
49. Zheng, C. et al. Cross-linking measurements of in vivo protein complex topologies. *Mol. Cell Proteomics* **10**, M110.006841 (2011).
50. Rozbesky, D. et al. Impact of chemical cross-linking on protein structure and function. *Anal. Chem.* **90**, 1104–1113 (2018).

51. Ding, Y. H. et al. Modeling protein excited-state structures from “over-length” chemical cross-links. *J. Biol. Chem.* **292**, 1187–1196 (2017).
52. Singh, P., Panchaud, A. & Goodlett, D. R. Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique. *Anal. Chem.* **82**, 2636–2642 (2010).
53. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).
54. Fischer, L. & Rappsilber, J. Quirks of error estimation in cross-linking/mass spectrometry. *Anal. Chem.* **89**, 3829–3833 (2017).
55. Trnka, M. J. et al. Matching cross-linked peptide spectra: only as good as the worse identification. *Mol. Cell Proteomics* **13**, 420–434 (2014).
56. Keller, A., Chavez, J. D. & Bruce, J. E. Increased sensitivity with automated validation of XL-MS cleavable peptide crosslinks. *Bioinformatics* **35**, 895–897 (2019).
57. Keller, A. et al. A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol. Syst. Biol.* **1**, 2005.0017 (2005).
58. Keller, A. et al. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392 (2002).
59. Shteynberg, D. et al. iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell Proteomics* **10**, M111.007690 (2011).
60. Kulak, N. A. et al. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324 (2014).
61. Liang, K. et al. Cryo-EM structure of human mitochondrial trifunctional protein. *Proc. Natl. Acad. Sci. USA* **115**, 7039–7044 (2018).
62. Jacobs, A. C. et al. AB5075, a highly virulent isolate of *Acinetobacter baumannii*, as a model strain for the evaluation of pathogenesis and antimicrobial treatments. *MBio* **5**, e01076-14 (2014).

### Acknowledgements

We thank members of the Bruce Lab for helpful discussions and suggestions. This work was supported by the following NIH grants: R01GM086688 to J.E.B., J.D.C., and A.K.; R01GM097112 to J.E.B. and J.P.M.; U19AI107775 to J.E.B., X.Z., and A.K.; R01HL110349 to J.E.B. and J.D.C.; and R01HL142628 to J.E.B. and J.D.C.

### Author contributions

J.D.C., J.P.M., M.M., X.Z., A.K., and J.E.B. contributed to the development of this protocol. J.D.C., J.P.M., and J.E.B. wrote and edited the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41596-019-0181-3>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to J.E.B.

**Peer review information:** *Nature Protocols* thanks Andrea Sinz and other anonymous reviewer(s) for their contribution to the peer review of this work.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 24 August 2018; Accepted: 18 April 2019;

Published online: 3 July 2019

### Related links

#### Key references using this protocol

Chavez, J. D., Schweppe, D. K., Eng, J. K., & Bruce, J. E. *Cell Chem. Biol.* **23**, 716–726 (2016): <https://doi.org/10.1016/j.chembiol.2016.05.012>

Wu, X. et al. *Nat. Commun.* **7**, 13414 (2016): <https://doi.org/10.1038/ncomms13414>

Mohr, J. P., Perumalla, P., Chavez, J. D., Eng, J. K., & Bruce, J. E. *Anal. Chem.* **90**, 6028–6034 (2018): <https://doi.org/10.1021/acs.analchem.7b04991>

Keller, A., Chavez, J. D., & Bruce, J. E. *Bioinformatics* **35**, 895–897 (2019): <https://doi.org/10.1093/bioinformatics/bty720>

Keller, A., Chavez, J. D., Eng, J. K., Thornton, Z., & Bruce, J. E. *J. Proteome Res.* **18**, 753–758 (2019): <https://doi.org/10.1021/acs.jproteome.8b00703>

#### Key data used in this protocol

Chavez, J. D. et al. *Cell Syst.* **6**, 136–141.e5 (2018): <https://doi.org/10.1016/j.cels.2017.10.017>

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Confirmed   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The exact sample size ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted<br><i>Give P values as exact values whenever suitable.</i>                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i> ), indicating how they were calculated   |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection	Thermo Xcalibur 2.1
Data analysis	<ul style="list-style-type: none"> <li>• ReAdW (<a href="https://sourceforge.net/projects/sashimi/files/ReAdW%20%28Xcalibur%20converter%29/">https://sourceforge.net/projects/sashimi/files/ReAdW%20%28Xcalibur%20converter%29/</a>) - converts .RAW files into .mzXML format</li> <li>• Mango (<a href="https://github.com/jpm369/mango">https://github.com/jpm369/mango</a>) - Identifies PIR mass relationships from mass spectrometry data</li> <li>• Comet 2018.01 or forward (<a href="http://comet-ms.sourceforge.net/">http://comet-ms.sourceforge.net/</a>) - Assigns peptide sequences to mass spectrometry data</li> <li>• XLinkProphet (<a href="https://github.com/brucelab/xlinkprophet">https://github.com/brucelab/xlinkprophet</a>) - Assigns probabilities and estimates false discovery rate for cross-linked peptide pair identifications from mass spectrometry data</li> <li>• Perl v5.24.0+ (<a href="https://www.perl.org/get.html">https://www.perl.org/get.html</a>) - Perl 5 is a highly capable, feature-rich programming language with over 30 years of development. Perl 5 runs on over 100 platforms from portables to mainframes and is suitable for both rapid prototyping and large scale development projects. Perl is needed for running XLinkProphet.</li> </ul>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. This is a protocol for carrying out chemical cross-linking with mass spectrometry experiments. A sample size of 1 is sufficient for discovery purposes.
Data exclusions	No data was excluded.
Replication	The described protocol has been used successfully to identify cross-linked peptide pairs from E. coli, HeLa cells, and mouse heart tissue. There are multiple published results demonstrating the repeatability of the protocol using biological replicates. For E. Coli see: doi: 10.1021/pr3011638, doi: 10.1021/acs.jproteome.6b00752, doi:10.1021/acs.analchem.7b04991. For HeLa see: doi:10.1074/mcp.M112.024497, doi:10.1038/ncomms8928, doi:10.1021/acs.analchem.7b04991, doi:10.1016/j.chembiol.2016.05.012. For mouse see: doi:10.1016/j.cels.2017.10.017
Randomization	No comparison between experimental groups is performed in the described protocol.
Blinding	Blinding is not necessary for investigators to follow the described protocol.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HeLa from ATCC
Authentication	ATCC Cell Line Authentication Service
Mycoplasma contamination	cell lines were not tested for mycoplasma contamination
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	None

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Mus musculus, strain C57BL/6NCrl, male and female, 12-14 weeks old
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve field-collected samples.



## Ethics oversight

All procedures involving animal use were performed with the approval of Institutional Animal Care and Use Committee of the University of Washington.

Note that full information on the approval of the study protocol must also be provided in the manuscript.