Rapid Mapping of Protein Interactions Using Tag-Transfer Photocrosslinkers

Jim E. Horne\textsuperscript{a}, Martin Walko\textsuperscript{a}, Antonio N. Calabrese\textsuperscript{a}, Mark A. Levenstein, David J. Brockwell, Nikil Kapur, Andrew J. Wilson,\textsuperscript{*} and Sheena E. Radford\textsuperscript{*}

Abstract: Analysing protein complexes by chemical crosslinking-mass spectrometry (XL-MS) is limited by the side-chain reactivities and sizes of available crosslinkers, their slow reaction rates, and difficulties in crosslink enrichment, especially for rare, transient or dynamic complexes. Here we describe two new XL reagents that incorporate a methanethiosulfonate (MTS) group to label a reactive cysteine introduced into the bait protein, and a residue-unbiased diazirine-based photoactivatable XL group to trap its interacting partner(s). Reductive removal of the bait transfers a thiol-containing fragment of the crosslinking reagent onto the target that can be alkylated and located by MS sequencing and exploited for enrichment, enabling the detection of low abundance cross-links. Using these reagents and a bespoke UV LED irradiation platform, we show that maximum crosslinking yield is achieved within 10 seconds. The utility of this “tag and transfer” approach is demonstrated using a well-defined peptide/protein regulatory interaction (BID\textsuperscript{MCL-1}), and the dynamic interaction interface of a chaperone/substrate complex (Skp/OmpA).

Chemical crosslinking-mass spectrometry (XL-MS) is a powerful component of the structural biology toolkit.\textsuperscript{[1]} XL-MS methods have been enhanced by new data analysis software,\textsuperscript{[2]} cleavable crosslinkers,\textsuperscript{[3]} strategies for crosslink enrichment,\textsuperscript{[4]} footprinting reagents,\textsuperscript{[5]} and structural modelling approaches.\textsuperscript{[6]} However, several challenges remain. For example, many XL-MS reagents have limited chemical reactivities (e.g. succinimide ester-based reagents)\textsuperscript{[7]} restricting the residues for which information can be obtained. Secondly, the low-abundance of crosslinked products often necessitates enrichment prior to MS.\textsuperscript{[8]} Finally, crosslinked peptide identification is difficult due to spectral complexity, poor fragmentation efficiency, and the increased search space associated with the large numbers of peptide combinations.\textsuperscript{[9]} Bulky crosslinking reagents may also perturb native interactions or create new aberrant interactions.\textsuperscript{[10]} Furthermore, many biological processes, for example, protein folding, binding and conformational changes, occur on short timescales. Consequently, crosslinks detected for a dynamic/non-equilibrium system report the average of multiple states or are dominated by the longest-lived state of all those populated. Photocrosslinkers such as diazirines can address this challenge and have been used previously to tag Cys residues in peptides.\textsuperscript{[11]} Diazirine-generated carbene react with proteins in ns,\textsuperscript{[11]} yet long irradiation times (minutes to hours)\textsuperscript{[10b,12]} are often required to generate acceptable crosslink yields due to the use of low intensity lamps.

Here we exploit a “tag and transfer” approach\textsuperscript{[13]} to develop two new XL-MS reagents and a workflow that enables crosslink identification for both well-defined and dynamic protein-protein interactions (PPIs). These heterobifunctional reagents comprise a methanethiosulfonate (MTS) group for specific attachment onto a single Cys residue introduced into the “bait” protein, creating a cleavable disulfide bond within the linker arm, and a diazirine that crosslinks to a “target” protein (Figure 1a,b, Figure S1,S2). Commercially available MTS-benzophenone-biotin tags have been used in photoinduced-crosslinking (PI-XL), but these bulky tags (752 Da) may perturb PPI interfaces.\textsuperscript{[14]} In the reagents described here, reduction of the disulfide bond-containing linker arm between the crosslinked proteins leaves a thiol tag on the target (87 or 204 Da in size) (Figure 1b, Figure S3). The modified residue can then be alkylated (e.g. with iodoacetamide [IAA]) and localised using MS protocols for identifying post-translational modifications (PTMs),\textsuperscript{[15]} or enriched before MS. We also describe the construction of a 365 nm UV LED irradiation platform (Figure 1c, Figure S4a,b) which enables crosslinking reactions to be completed in just 10 seconds with marginal heating. We exemplify this methodology by mapping two PPIs, each with a different binding mode: a complex between BID\textsuperscript{MCL-1} (K\textsubscript{D} = 50 ± 20 nm),\textsuperscript{[15]} where the tight binding affinity is...
mediated by several key residues\cite{15a,16} and the dynamic chaperone/substrate complex Skp/OmpA (\(K_D = 22 \pm 16\) nm),\cite{17} in which the binding interface involves many rapidly interconverting interactions\cite{17,18}.

MTS-diazirine and MTS-trifluoromethyl phenyl diazirine (MTS-TFMD) (Figure 1a, Figure S1) were chosen as photo-activatable groups due to the superior performance of diazirines in comparative PI-XL studies\cite{13c,19}. Their small size, and rapid, indiscriminate reactivity (Figure S2).\cite{9,11} Both diazirine- and TFMD-containing crosslinkers were synthesised as the photochemistry of TFMD leads to higher cross-linking efficiency, but it is bulkier.\cite{13b,18} In the bait, a unique solvent exposed Cys for crosslinker conjugation is introduced by mutagenesis or synthesis. In proteins/peptides which lack Cys, or contain buried Cys, this is straightforward. When a solvent exposed Cys is already present in the bait this can be exploited, or substituted with Ala or Ser, and a new Cys introduced in a location of interest. Knowing the location of the Cys on one partner in the PPI reduces the MS/MS search space from \(n_{\text{peptide}}\) to \(n_{\text{target}}\) (where \(n_{\text{peptide}}\) is the number of crosslinkable residues in the bait and target—for diazirines, this is every residue in the proteins—and \(n_{\text{target}}\) is the number of residues in the target). Here, we introduce a unique Cys by mutagenesis into the E. coli outer membrane protein OmpA, while for the BID\(_{80-102}\) peptide, which comprises the binding domain of the pro-apoptotic protein BID, Cys was introduced via solid-phase peptide synthesis. Importantly, the reactivity of the photoactivated diazirine places no restrictions on the amino acids in the target protein that can be detected once a complex is formed.

To enhance diazirine photoactivation efficiency we constructed a 365 nm LED lamp (Figure 1c, Figure S4a,b) with a nominal flux density of 0.2 μmol cm\(^{-2}\). We designed devices to irradiate samples in 0.2 mL to 1.5 mL tubes, or in acrylic chip\cite{21} with 33 μl sample wells (Figure 1c, Figure S5). Using this LED platform, maximal XL yields were achieved in 10 seconds, with an associated heating of \(< 2\) °C (Figure 2a,b, Figure S4c). By contrast, a 6 W Hg-Xe lamp achieved maximal yields after ca. 20 minutes with heating of 10 °C (Figure 2a,b, Figure S4c), consistent with previous reports.\cite{10,12c,19} The UV LED design thus improves the reaction rate 130-fold, and decreases heating 6-fold.

The human apoptotic regulatory pair BID/MCL-1 is a target for cancer drug discovery.\cite{22} The BH3 binding domain of BID (BID\(_{80-102}\)) adopts a helix on binding to a surface groove on MCL-1 (Figure 3a)\cite{15b,16}. Five single Cys variants of BID\(^{80-102}\) (Figure 3a) were synthesised (see Supporting Information), each with a unique Cys residue in a different position. Subsequently, the peptides were tagged with MTS-diazirine or MTS-TFMD. The modified peptides bound MCL-1 with high affinity, although those with the tagged residues in the centre of BID\(_{80-102}\) (B86C and V93C) had a reduced EC\textsubscript{50} (Figure 3b, Table S1). Crosslinking of the BID\(_{80-102}\) peptides with MCL-1 was achieved by UV LED irradiation. Crosslinked complexes were detected by SDS-PAGE (Figure 3c). The absolute crosslinking efficiencies varied from 9–40% for MTS-diazirine, with reduced efficiency for lower affinity variants (Table S1). Absolute cross-linking efficiencies were higher for the MTS-TFMD tagged peptides (26–53%) (Table S1), as expected given the greater yield of TFMD-derived carbenes.\cite{20} Gel bands of the crosslinked complexes were excised and the linker arm cleaved by reduction of the disulfide (Figure 1b) releasing MCL-1 bearing the transferred thiol-containing tag at the interaction sites. The free thiol in each tag was capped by alkylation with IAA and the protein digested in-gel with trypsin (see Supporting Information). Peptides modified with the transferred tag were identified by LC-MS/MS. This allowed BID\(_{80-102}\) to be mapped.
into the MCL-1 binding groove unambiguously, in agreement with the known “bind-and-fold” interaction\([15,16]\) and NMR/X-ray structures of the complex (Figure 3d, Tables S2 and S3, Figures S6–S8).\([15,16]\) Both crosslinkers gave similar results (Figure S6), demonstrating that the bulkier TFMD-diazirine can be used even when sidechain interdigitation drives association. Crosslinked positions on MCL-1 were quantified by MS/MS from each Cys residue introduced into BID\(_{80–102}\) labelled with MTS-diazirine or MTS-TFMD. This revealed that crosslinking yields were greater when the Ca–Ca Euclidian distance of the residues in the complex was <10 Å, but crosslinks could be detected for distances up to 15 Å (Figures S9,S10).

We next tested the ability of our workflow and tag transfer reagents to study a protein complex stabilised by transient interactions. Chaperone-client binding often involves a dynamic interaction of chaperones with an unfolded/partially folded client protein to aid folding or prevent aggregation.\([17,18]\) These interfaces are challenging to map using crosslinking since the interactions can be dynamic and diffuse. Here, we used the interaction between the periplasmic chaperone Skp from \textit{E. coli}, and a β-barrel outer membrane protein substrate, OmpA, as a model for this type of PPI.\([19]\) Skp, a homotrimer, has an jellyfish-like structure resulting in a cage in which substrates are sequestered (Figure 4a).\([20,21]\) Skp binds its OMP substrates with n affinity,\([22]\) similar to that of the well-defined BID\(_{80–102}/\text{MCL-1}\) complex.\([23]\) However, Skp-bound OmpA is in a “fluid-globule” state that forms many weak and transient interactions in the Skp cage.\([24,25]\)

Two single Cys variants of OmpA were generated, W7C and T144C, located in the β-barrel and a surface exposed loop in folded OmpA, respectively.\([26]\) MTS-diazirine or MTS-TFMD conjugation of both variants was efficient (Figure S11), and did not prevent folding (Figure S12). The Skp/OmpA complex was assembled by dilution of urea-denatured OmpA into a Skp-containing solution, and PI-XL was then performed. The cross-linked Skp/OmpA complex was detected by non-reducing SDS-PAGE as a single band with an apparent mass of approximately 70 kDa (Figure 4b). Four non-overlapping positions from OmpA-(T144C) to residues within Skp’s “cage” (Figure 4d (top); Figure S13) could be identified by in-gel digestion of the crosslinked proteins followed by LC-MS/MS (Figure 5, Method 1). This result was surprising since it is inconsistent with previous studies which have shown that OmpA tumbles dynamically on a sub-ms timescale within Skp.\([27]\) We reasoned that the lack of modified sites resulted from the relatively low abundance of these peptides in the Skp-OmpA complex, rather than reflecting a specific interaction surface involving these four sites. Given that each Skp trimer binds a single OmpA, and each OmpA contains only a single crosslinker, then only one modified peptide would result from each 70 kDa Skp/OmpA complex.

To test this conclusion we developed a protocol to enrich chemically modified Skp, removing background peptides arising from OmpA and unmodified Skp (Figure 5). We used thiopropyl Sepharose 6B beads to purify Skp monomers containing the thiol fragment obtained from the tag-transfer reaction and the Cys-containing OmpA bait. The thiol-containing proteins were then eluted from the resin, separated by SDS-PAGE and tryspinised in-gel (Method 2). Alternatively, trypsin digestion was performed on-bead\([28]\) and unbound peptides removed by washing the resin prior to elution (Method 3). Using Method 2 the number of modified peptides detected increased >2-fold (from 6 to 14), whilst Method 3 yielded a further 2.5-fold increase in the number of modified peptides identified (from 14 to 35) (Figure 4d, Figure S13), a 6-fold increase in detection over Method 1. No significant sidechain bias was observed for either crosslinker (Figure S14), consistent with the reactivity of the diazirine-derived carbenes.\([9,16]\) Similar crosslinking sites were observed for OmpA(W7C) and OmpA(T144C) using both crosslinkers (Figures S13,S14). Using these enrichment protocols modifications on Skp were identified all around the internal cavity (Figure 4c), consistent with OmpA tumbling randomly in Skp.\([10]\) The ability to enrich crosslinked peptides...
using our tag-transfer protocol, combined with the promiscuous reactivity of diazirines, exemplifies the power of the technique to monitor even the most dynamic of protein interfaces (Figure 4c & d).

In summary, we have demonstrated that Cys-containing variants of a bait protein conjugated with MTS-diazirine or MTS-TFMD-based tag-transfer crosslinkers can map PPIs in both well-defined and dynamic interfaces. Since the location of the crosslink in the bait protein is known and only the transferred tag is considered in downstream analysis, the background from bait peptides is removed and the search space for crosslinked products is reduced. These features are particularly important as target protein size increases. The workflow described is simple to implement, only requiring the appropriate crosslinking reagents, LC-MS/MS and proteomics software for mapping PTMs. Enrichment and digestion of modified peptides enables low protein concentrations to be used, minimising the possibility of aggregation or other aberrant interactions. Additionally, the custom UV LED platform enables PI-XL on a 10 second timescale, not possible with arc-based lamps, and only previously achieved using pulsed lasers in solution or in the gas phase.\(^5,10^a\) The low cost (≈$300) and simplicity of our UV LED system makes these timescales accessible to any researcher. Our workflow thus opens the door to time-resolved XL on the second timescale without the need for expensive lasers and enables the study of conformational changes within dynamic protein complexes versus time.

**Acknowledgements**

We thank Nasir Khan and Rachel George for technical assistance; Dr. Thomas Edwards and Pallavi Ramsahye for MCL-1; Dr. Bob Schiffrin for the collapsed model of tOmpA; and Patrick Mason for use of an IR camera. His-tagged Skp plasmid was kindly provided by S. Hiller (University of Basel, Switzerland). Full-length OmpA plasmid was kindly provided by K. Fleming (John Hopkins University, USA). The BBSRC (BB/M011151/1, BB/K000659/1, BB/N007603/1 and BB/M012573/1, BB/P00037/1), EPSRC (EP/N035267/1 and EP/N013573/1) and Wellcome Trust (208385/Z/17/Z) are acknowledged for their support. M.A.L. is supported by a Leeds International Research Scholarship. N.K. acknowledges the support of RAEng and GSK. S.E.R. acknowledges...
Conflict of interest

The authors declare no conflict of interest.

Keywords: chemical crosslinking · diazo compounds · mass spectrometry · photoaffinity labeling · protein–protein interactions

How to cite: Angew. Chem. Int. Ed. 2018, 57, 16888–16902

Angew. Chem. 2018, 130, 16930–16934

[1] a) M. Schneider, A. Belsom, J. Rappisilber, Trends Biochem. Sci. 2018, 43, 157–169; b) A. N. Calabrese, S. E. Radford, Methods 2018, 147, 187–205; c) J. Gubbens, E. Ruijter, L. E. V. de Ruijs, J. M. A. Damen, B. de Kruijff, M. Slipper, D. T. S. Rijkers, R. M. J. Liskamp, A. L. P. M. de Kroon, Chem. Sci. 2009, 16, 3–14; d) O. Pavlova, J. H. Peterson, R. Ieva, H. D. Bernstein, Proc. Natl. Acad. Sci. USA 2013, 110, E938–947; e) C. R. Braun, J. M. U. V. de Fays, G. H. Bird, S. P. Gygi, L. D. Walensky, Chem. Biol. 2010, 17, 1325–1333; f) A. Sinz, Angew. Chem. Int. Ed. 2018, 57, 6390–6396; Angew. Chem. 2018, 130, 6498–6504.

[2] S. Yilmaz, G. A. Shiferaw, J. Rayo, A. Economou, L. Martinus, E. Vandenbriele, Mass Spectrom. Rev. 2018, 37, 738–749.

[3] a) C. Iacobucci, C. Piotrowski, A. Rehkamp, C. H. Ihling, A. Sinz, J. Am. Soc. Mass Spectrom. 2018, 29, 0; b) M. Q. Müller, F. Dreickoer, C. H. Ihling, M. Schäfer, A. Sinz, Anal. Chem. 2010, 82, 6958–6968; c) Y. Yang, H. Song, D. He, S. Zhang, S. Dai, X. Xie, S. Lin, Z. Hao, H. Zheng, P. R. Chen, Nat. Protoc. 2017, 12, 2147–2168; d) C. Iacobucci, M. Gotze, C. Piotrowski, C. Arlt, A. Rehkamp, C. Ihling, C. Hage, A. Sinz, Anal. Chem. 2018, 90, 2805–2809.

[4] D. Tan, Q. Li, M. J. Zhang, C. Liu, C. Ma, P. Zhang, Y. H. Ding, S. B. Fan, L. Tao, B. Yang, X. Li, S. Ma, J. Liu, B. Feng, X. Liu, H. W. Wang, S. M. He, N. Gao, K. Ye, M. Q. Dong, X. Lei, eLife 2016, 5, e12509.

[5] a) L. Manzi, A. S. Barrow, D. Scott, R. Layfield, T. G. Wright, J. E. Moses, N. J. Common, Nat. Commun. 2016, 7, 13288; b) L. Manzi, A. S. Barrow, J. T. S. Hopper, R. Kaminissa, C. Kleinthous, C. V. Robinson, J. E. Moses, N. J. Oldham, Angew. Chem. Int. Ed. 2017, 56, 14873–14877; Angew. Chem. 2017, 129, 15069–15073.

[6] a) A. Politis, C. Schmidt, J. Proteomics 2018, 175, 34–41; b) M. Ferber, J. Kosinski, A. Ori, U. J. Rashid, M. Moreno-Morcillo, B. Simon, G. Bouvier, P. R. Batista, C. W. Muller, M. Beck, M. Nilges, Nat. Methods 2016, 13, 515–520.

[7] a) S. Kalkhof, A. Sinz, Anal. Bioanal. Chem. 2008, 392, 305–312; b) S. Mäder, C. Bich, D. Toutouh, R. Zenobi, J. Mass Spectrom. 2009, 44, 612–609.

[8] a) R. Fritzsche, C. H. Ihling, M. Gotze, A. Sinz, Rapid Commun. Mass Spectrom. 2012, 26, 653–658; b) R. M. Kaake, X. Wang, A. Burke, C. Yu, W. Kandur, Y. Yang, E. J. Novitsky, T. Second, J. Duan, A. Kao, S. Guan, D. Velluci, S. D. Rychnovský, L. Huang, Mol. Cell. Proteomics 2014, 13, 3533–3543; c) A. Leitner, R. Reischl, T. Walzthoeni, F. Herzog, S. Bohn, F. Forster, R. Aebersold, Mol. Cell. Proteomics 2012, 11, M111.014126; d) X. Tang, J. E. Bruce, Mol. Biosyst. 2010, 6, 939–947.

Manuscript received: August 8, 2018
Revised manuscript received: October 30, 2018
Accepted manuscript online: November 4, 2018
Version of record online: November 21, 2018