A Click-Chemistry-Based Enrichable Crosslinker for Structural and Protein Interaction Analysis by Mass Spectrometry

Michael Stadlmeier+,[a, b] Leander Simon Runtsch+,[a] Filipp Streshnev,[a] Martin Wühr,[b] and Thomas Carell*[a]

Mass spectrometry is the method of choice for the characterisation of proteomes. Most proteins operate in protein complexes, in which their close association modulates their function. However, with standard MS analysis, information on protein–protein interactions is lost and no structural information is retained. To gain structural and interactome data, new crosslinking reagents are needed that freeze inter- and intramolecular interactions. Herein, the development of a new reagent, which has several features that enable highly sensitive crosslinking MS, is reported. The reagent enables enrichment of crosslinked peptides from the majority of background peptides to facilitate efficient detection of low-abundant crosslinked peptides. Due to the special cleavable properties, the reagent can be used for MS^2 and potentially for MS^3 experiments. Thus, the new crosslinking reagent, in combination with high-end MS, should enable sensitive analysis of interactomes, which will help researchers to obtain important insights into cellular states in health and diseases.

Proteins need to interact with other proteins to form functional complexes. In many cases, protein function inside cells cannot be understood without information about the protein structure and knowledge of in which complex the protein is situated.[1,2] This is, for example, of paramount importance for proteins that modulate epigenetic information on DNA. Almost all of these chromatin-modifying proteins require intensive interaction with metabolic enzymes, which provide the cofactors needed for histone acetylation, deacetylation or methylation

and demethylation.[3–5] This showcases the need to study the complex environment of a given protein to analyse its function and activity state. Protein crosslinking, in combination with analysis by mass spectrometry (XL-MS), is ideally suited as a method to gain information about protein structure and the composition of protein complexes.[6–9] For XL-MS, specialised chemical reagents, crosslinkers, are required, which are able to covalently connect protein residues that are in close proximity, for example, interacting in a complex.[10, 11] Characterisation of the crosslinked peptides by means of mass spectrometry, however, poses a formidable challenge for two reasons. Firstly, crosslink identification has to be achieved by analysing fragment ions of not only one, but two, connected peptides, thereby massively complicating MS^2 spectra. Secondly, the crosslinked peptide species have only a very low abundance and are part of a peptide mixture that is overwhelmingly dominated by non-crosslinked peptides. In many cases, this leads to dramatic information loss that hampers accurate crosslink identification, and therefore, interactome analysis. A few crosslinking reagents were developed, which introduced MS-cleavable groups, and the possibility for XL enrichment to tackle these problems.[12–18]

Herein, we report the development of a new crosslinking reagent that is enrichable and has cleavage properties which allow accurate crosslink identification. The designed cliXlink (**1**) is depicted in Figure 1. The cell-permeable reagent (Figure 1 A and Figure S1 in the Supporting Information) features two succinimidyl ester units that can react with nucleophiles in proteins and are spaced roughly 9 Å apart; thus enabling the fixation of short distances to gain structural information on proteins and to capture proteins in close proximity to each other. Efficient MS cleavability is ensured by the well-established sulfoxide group, which can be cleaved under low-energy collision-induced dissociation (CID) conditions prior to peptide fragmentation. In addition, an alkyne unit allows an enrichment moiety, for example, biotin, to be attached to the crosslinked sites with the help of CuAAC reaction.[19]

Application of the crosslinker can follow the workflow depicted in Figure 1B. It involves the addition of reagent **1** to a complex proteome, fixing the native state distance information of peptides in close proximity and preserving them throughout the further workflow for MS analysis. Attachment of the affinity group for enrichment is performed after crosslinking to avoid interference from bulky enrichment groups. By modification of the crosslinked peptides at the protein level, excess small-molecule reagents can readily be removed by acetone precipita-

[a] Dr. M. Stadlmeier,+ L. S. Runtsch,+ F. Streshnev, Prof. Dr. T. Carell
Department of Chemistry, Ludwig-Maximilians-Universität München
Butenandtstrasse 5–13, 81377 Munich (Germany)
E-mail: thomas.carell@lmu.de

[b] Dr. M. Stadlmeier,+ Prof. Dr. M. Wühr
Department of Molecular Biology and
the Lewis-Sigler Institute for Integrative Genomics, Princeton University
Washington Road, Princeton, NJ 08544 (USA)

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201900611.
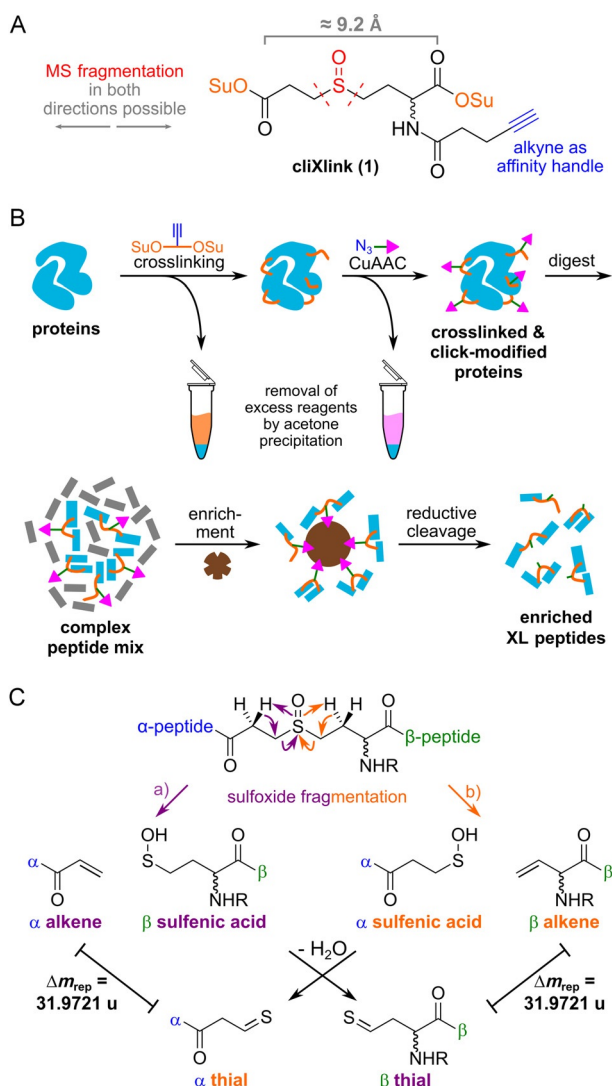
Figure 1. A) Structure of **1**. Su: succinimidyl. B) Workflow for XL-MS experiments. After the addition of **1**, protein sites in close proximity are covalently linked. The generated crosslinks can be functionalised with a copper-catalysed Huisgen reaction (CuAAC). Excess reagents are removed by acetone precipitations. After enzymatic digestion and enrichment on magnetic streptavidin beads, crosslinked peptides were cleaved off under mild conditions and subsequently analysed by means of LC-MS[2]. C) Fragmentation pathways of the crosslinked peptides, forming two peptide pairs with a characteristic Δm_{rep} of 31.9721 u.

to a clean separation of the peptides (α, β) and generates two fragments for each peptide: an alkene and a sulfenic acid fragment; the latter forms a thial upon water loss. This provides two mass pairs with a characteristic $\Delta m/z \approx 32$. Importantly, we observed that fragmentation pathway a (Figure 1 C) predominated. Therefore, for the α-peptide the alkene fragment, and for the β-peptide the thial fragment, are the main cleavage products. Due to the asymmetric cleavage properties, most of the signal intensity is retained on one fragment per peptide, which should provide excellent sensitivity in MS[3] experiments.

The synthesis of reagent **1** is straightforward (Scheme 1). The starting point is the oxidised disulfide dimer of homocystein **2**, which is treated with 4-pentynoic acid to give disulfide
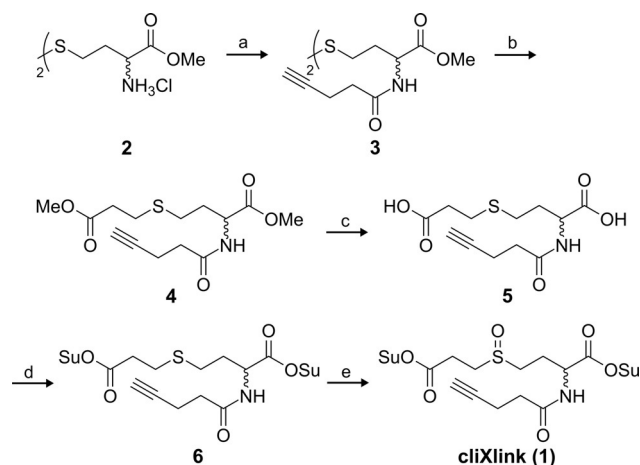
Scheme 1. Synthesis of cliXlink (**1**). a) 4-Pentynoic acid, *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl), 1-hydroxybenzotriazole hydrate (HOBt·H$_2$O), NEt$_3$, DMF, RT, 15 h, 71%; b) over two steps: 1) tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl), NaHCO$_3$, DMF, H$_2$O, RT, 3 h; 2) methyl 3-bromopropionate, 45 °C, 44 h, 55%; c) LiOH, THF, H$_2$O, 0 °C RT, overnight (o/n), 91%; d) *N*-hydroxysuccinimide (NHS), pyridine, trifluoroacetic anhydride, MeCN, 0 °C RT, o/n, 78%; e) *meta*-chloroperoxybenzoic acid (*m*CPBA), AcOEt, RT, 30 min, 57%.

3. Reductive cleavage of the disulfide and alkylation of the generated thiols with methyl 3-bromopropionate generates sulfide compound **4**. Saponification of both methyl esters to **5** and conversion of **5** into activated bissuccinimidyl ester **6**, followed by oxidation of the thioether to the sulfoxide, gives reagent **1** in a total yield of 16%. Of particular importance is that the reagent is very pure and the reactive ester units are in place at both sides. Partial hydrolysis needs to be avoided. This was ensured by a final precipitation purification. Reagent **1** was dissolved in a mixture of ethyl acetate and dichloromethane and precipitated upon addition of hexanes.

We next investigated the MS properties of **1**. To this end, we added **1** to the commonly used model protein, bovine serum albumin (BSA). We followed the workflow depicted in Figure 1 B without performing the enrichment step. In short, after the addition of **1** to the protein solution and reaction for 1 h at room temperature and physiological pH, we precipitated the protein with acetone, resuspended it in buffer (see the

tion. This is followed by enzymatic digestion of the proteins. The crosslinked peptides are in this way labelled, for example, with biotin, which allows for their enrichment. Afterwards, they are analysed by means of mass spectrometry. Because the individual peptide masses in a crosslink are not immediately accessible, crosslink assignment is difficult, especially in complex samples. This requires the use of MS-cleavable reagents, which facilitate MS[2] identification by forming specific fragment ions.[20,21] At best, the reagent should also enable MS[3] experiments, which requires that the crosslinker is cleaved before the peptides. This allows peptides to be separated for individual MS[3]-based identification. Our reagent **1** contains β-hydrogen atoms on both sides of the sulfoxide, so that fragmentation occurs in two directions, as depicted in Figure 1 C. This leads

Supporting Information) and digested the protein subsequently with a mixture of trypsin and Lys-C.

The obtained peptide mixture was desalted with C18 Tip columns and analysed by means of HPLC-MS^2 (Figure 2). Figure 2A shows two crosslinked BSA peptides ($\alpha+\beta$) as an example. After determining the exact mass of the intact crosslink (*m/z* 677.1; Figure 2B), we performed both CID and HCD fragmentation on the precursor to evaluate the fragmentation properties (Figure 2C). More selective CID fragmentation (resonance excitation of the crosslink ion; Figure 2C, top spectrum) at a low normalised collision energy of 25% provides, as expected, only a small set of intense signals. Two prominent signal pairs, with Δ*m/z* of 32 (Δm_{rep}), are obtained due to crosslinker fragmentation, which results in a thial (α-/β-thial) and alkene (α-/β-alkene) fragment for each peptide. As mentioned before, fragmentation pathway a (Figure 1C) is preferred, leading to an asymmetric intensity distribution. The dominant formation of the expected intact, separated peptides consequently makes the reagent amenable for more sophisticated MS^3 experiments.

For our study, we used HCD fragmentation. This method provides simultaneous crosslinker cleavage and formation of the peptide fragments required for identification (Figure 2C, bottom spectrum). To help with crosslink/peptide recognition, it is important that HCD fragmentation still provides the alkene and thial fragments. The HCD data, therefore, allow the identification of peptides by MS^2.

Using this method, we analysed the data with the freely available MeroX software, strictly filtering all crosslink identifications (score > 50, false discovery rate (FDR) 1%) and requiring the presence of at least three out of four ions from the two characteristic mass pairs.[23, 24] Without exploiting the possibility for CuAAC-based enrichment, we were able to identify 61 unique BSA crosslinks in a single measurement (Figure 2D).[25] This result is representative for measurements conducted with purified BSA in our laboratory. Additionally, we depicted the crosslinks found in the BSA crystal stucture and noted that the majority of measured distances were reasonable. Importantly, the crosslinks show a mean C$_\alpha$–C$_\alpha$ distance between the crosslinked amino acids of 22.1 Å and a distance distribution that is in perfect agreement of what is expected for crosslinking with a reagent that spaces the active esters by about 9 Å (Figures 2E and S2)[26] and taking into account side-chain lengths and protein molecular dynamics.[27] Crosslinking occurred mostly between two Lys residues (49%), but we also detected Lys–Thr (30%), Lys–Ser (13%) and Lys–Tyr (8%) connections, in agreement with the reactivity of the succinimidyl esters (Figure 2F).[28]

This success allowed us to validate the new crosslinker in a more complex environment. We particularly wanted to show the additional value of the CuAAC-based enrichment possibility. For this experiment, we again crosslinked the BSA protein (10 µg), precipitated it with acetone, redissolved the crosslinked protein and subsequently performed the click reaction (Figure S3A) by adding **7** (Figure 3A) and CuSO$_4$/tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) followed by reduction of Cu^{II} to Cu^{I} upon addition of sodium ascorbate. After 1 h at
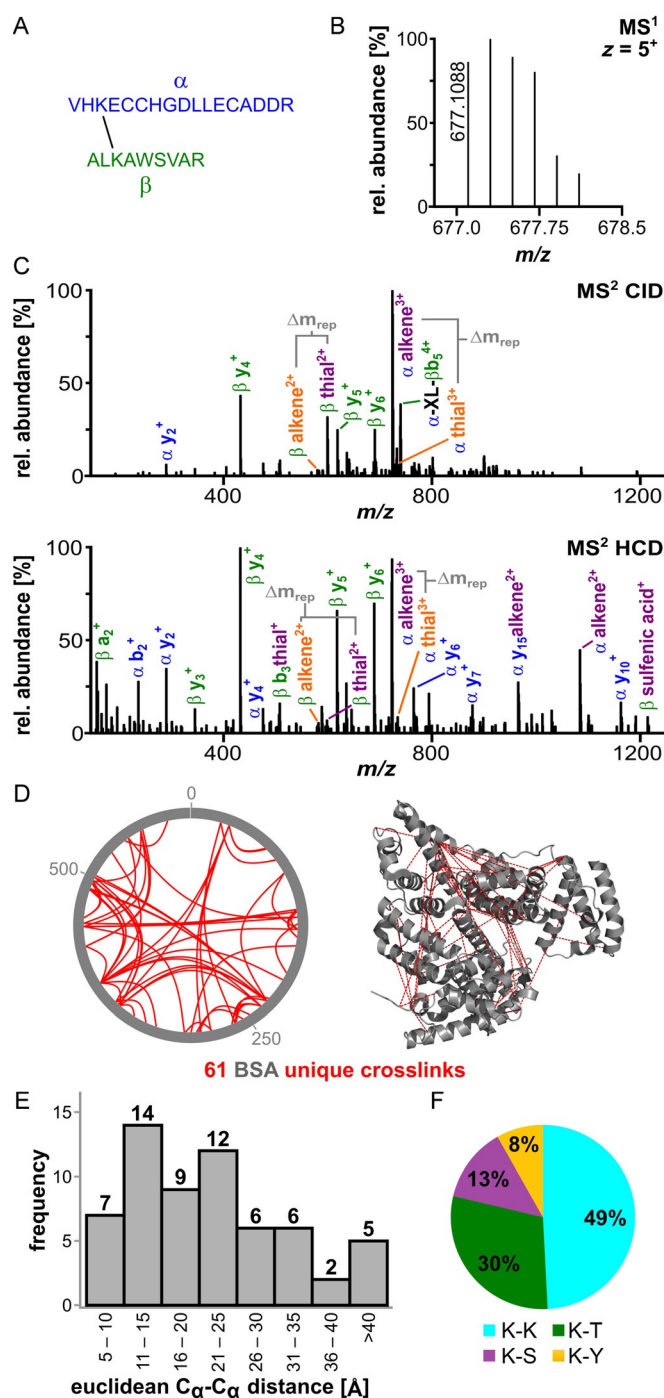

Figure 2. Analysis of crosslinked BSA prior to enrichment. A) Sequences of a representative crosslink within BSA: peptide α: VHKECCHGDLLECADDR (IAA-modified on Cys) and β: ALKAWSVAR. B) The isotope envelope of the 5+ precursor is shown. C) Fragmentation of the crosslinked peptides under CID (25%, top spectrum) and stepped higher energy collisional dissociation (HCD) conditions (25/30/32%, bottom spectrum) showing the product ions of sulfoxide cleavage. The different fragmentation pathways are shown in orange and purple (Figure 1). D) MS^2 identification of BSA crosslinking sites prior to enrichment. The xVis representation shows the positions of crosslinked amino acids as red lines (left). The corresponding C$_\alpha$–C$_\alpha$ distances are depicted in the BSA crystal structure (PDB ID: 4F5S[22]). E) Histogram of the measured Euclidean C$_\alpha$–C$_\alpha$ distances in the crystal structure; the majority are below 40 Å, with a mean measured distance of 22.1 Å. F) Distribution of the identified crosslinking sites between Lys–Lys, Lys–Thr, Lys–Ser and Lys–Tyr.

A

biotin-azide 7

B

HEK peptide background
435 µg

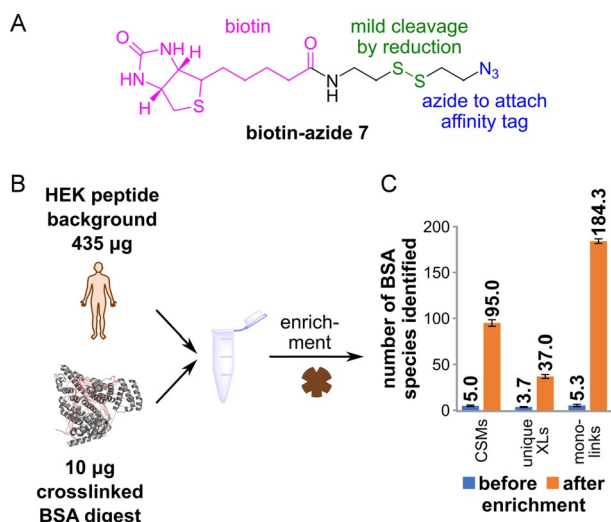10 µg crosslinked BSA digest

enrich-ment

C

Figure 3. Enrichment of crosslinked BSA from a complex sample. A) Structure of **7**, with the biotin group (pink), the disulfide bond (green) and the azide moiety (blue). Upon reduction of the disulfide, the biotin moiety is removed from the crosslink. B) Depiction of the spike-in experiment. 10 µg of crosslinked and CuAAC-modified BSA digest were added to 435 µg of HEK digest as a complex background. Magnetic streptavidin beads were used to enrich the crosslinked peptides. C) Results of the HPLC-MS^2 analysis, showing the effect of enrichment. The experiment was performed in technical triplicates; numbers show the mean value and the standard deviation is indicated by the error bar.

room temperature, we performed another acetone precipitation to remove excess biotin azide. We then added trypsin and Lys-C for digestion.

These peptides we next combined with a protein digest obtained from a HEK cell extract (435 µg of protein; Figure 3B). This creates a massive background of non-crosslinked peptides. If we now analysed the crosslinks within this large background (Figure 3C), we identified only a very small number of crosslink spectral matches (mean CSMs = 5.0), unique crosslinked sites (mean unique XLs = 3.7) as well as monolinks (mean = 5.3), in which one succinimidyl ester hydrolysed before reaction with the protein. However, if we performed the enrichment with streptavidin-coated magnetic beads, the number of crosslink identifications improved dramatically. Following incubation of the peptide mixture with the magnetic particles; extensive washing (6×) of the beads with buffer (see the Supporting Information) and mild, reductive cleavage of the disulfide to liberate the "captured" crosslinks and thereby removing the biotin moiety (Figure S3B), we now, on average, detected 95.0 CSMs and 37.0 unique XLs together with 184.3 monolinks. The number of identified unique XLs was lower than that for the purified BSA (Figure 2D), which could be explained by inefficiences in the CuAAC reaction or interference of the non-crosslinked complex background. However, this result shows that our new crosslinker, **1**, is able to enrich crosslinked peptides in a vast excess of unmodified peptides; thus facilitating the analysis of complex samples with low-abundant crosslinks. Originally, we were sceptical about the asymmetric structure of **1**. The data, however, show that, despite this, a large number of crosslinks are identified.

In summary, our new crosslinker **1** combines the following advantages: firstly, the size of the reagent is ideally suited to gain valuable distance information for structural proteomics. In addition, the molecule is cell permeable and the alkyne unit does not cause major interference with the crosslinking reaction. Furthermore, the robust, efficient sulfoxide fragmentation simplifies crosslink identification. The possibility of functionalising peptides crosslinked with **1** by click chemistry allows diverse modifications and enrichment strategies that enable analysis of low-abundant crosslinks to be employed. To conclude, our reagent **1** now paves the way for complex interactome analysis by using MS^2 and MS^3 experiments.

Conflict of Interest

The authors declare no conflict of interest.

[1] I. N. Berezovsky, E. Guarnera, Z. Zheng, B. Eisenhaber, F. Eisenhaber, *Curr. Opin. Struct. Biol.* **2017**, *42*, 67–74.

[2] H. Currinn, T. Wassmer, *Biochem. Soc. Trans.* **2016**, *44*, 185–190.

[3] S. Allard, R. T. Utley, J. Savard, A. Clarke, P. Grant, C. J. Brandl, L. Pillus, J. L. Workman, J. Côté, *EMBO J.* **1999**, *18*, 5108–5119.

[4] Y. Xue, J. Wong, G. T. Moreno, M. K. Young, J. Côté, W. Wang, *Mol. Cell* **1998**, *2*, 851–861.

[5] M. Sauvageau, G. Sauvageau, *PLoS Biol.* **2008**, *6*, e113.

[6] A. Leitner, T. Walzthoeni, A. Kahraman, F. Herzog, O. Rinner, M. Beck, R. Aebersold, *Mol. Cell. Proteomics* **2010**, *9*, 1634–1649.

[7] A. Sinz, *Angew. Chem. Int. Ed.* **2018**, *57*, 6390–6396; *Angew. Chem.* **2018**, *130*, 6498–6504.

[8] J. Rappsilber, *J. Struct. Biol.* **2011**, *173*, 530–540.

[9] J. D. Chavez, J. E. Bruce, *Curr. Opin. Chem. Biol.* **2019**, *48*, 8–18.

[10] F. J. O'Reilly, J. Rappsilber, *Nat. Struct. Mol. Biol.* **2018**, *25*, 1000–1008.

[11] A. Leitner, M. Faini, F. Stengel, R. Aebersold, *Trends Biochem. Sci.* **2016**, *41*, 20–32.

[12] A. Kao, C.-l. Chiu, D. Vellucci, Y. Yang, V. R. Patel, S. Guan, A. Randall, P. Baldi, S. D. Rychnovsky, L. Huang, *Mol. Cell. Proteomics* **2011**, *10*, M110.002212.

[13] M. Q. Müller, F. Dreiocker, C. H. Ihling, M. Schäfer, A. Sinz, *Anal. Chem.* **2010**, *82*, 6958 – 6968.

[14] A. M. Burke, W. Kandur, E. J. Novitsky, R. M. Kaake, C. Yu, A. Kao, D. Vellucci, L. Huang, S. D. Rychnovsky, *Org. Biomol. Chem.* **2015**, *13*, 5030 – 5037.

[15] X. Tang, J. E. Bruce, *Mol. Biosyst.* **2010**, *6*, 939 – 947.

[16] R. Huang, W. Zhu, Y. Wu, J. Chen, J. Yu, B. Jiang, H. Chen, W. Chen, *Chem. Sci.* **2019**, *10*, 6443 – 6447.

[17] M. Rey, M. Dupré, I. Lopez-Neira, M. Duchateau, J. Chamot-Rooke, *Anal. Chem.* **2018**, *90*, 10707 – 10714.

[18] B. Steigenberger, R. J. Pieters, A. J. R. Heck, R. A. Scheltema, *ACS Cent. Sci.* **2019**, *5*, 1514 – 1522.

[19] M. Ehrlich, M. J. Gattner, B. Viverge, J. Bretzler, D. Eisen, M. Stadlmeier, M. Vrabel, T. Carell, *Chem. Eur. J.* **2015**, *21*, 7701 – 7704.

[20] F. Liu, P. Lössl, R. Scheltema, R. Viner, A. J. R. Heck, *Nat. Commun.* **2017**, *8*, 15473.

[21] O. Klykov, B. Steigenberger, S. Pektaş, D. Fasci, A. J. R. Heck, R. A. Scheltema, *Nat. Protoc.* **2018**, *13*, 2964 – 2990.

[22] A. Bujacz, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 1278 – 1289.

[23] M. Götze, J. Pettelkau, R. Fritzsche, C. H. Ihling, M. Schäfer, A. Sinz, *J. Am. Soc. Mass. Spectrom.* **2015**, *26*, 83 – 97.

[24] M. Götze, C. Iacobucci, C. H. Ihling, A. Sinz, *Anal. Chem.* **2019**, *91*, 10236 – 10244.

[25] M. Grimm, T. Zimniak, A. Kahraman, F. Herzog, *Nucleic Acids Res.* **2015**, *43*, W362 – W369.

[26] E. D. Merkley, S. Rysavy, A. Kahraman, R. P. Hafen, V. Daggett, J. N. Adkins, *Protein Sci.* **2014**, *23*, 747 – 759.

[27] Z. Orbán-Németh, R. Beveridge, D. M. Hollenstein, E. Rampler, T. Stranzl, O. Hudecz, J. Doblmann, P. Schlögelhofer, K. Mechtler, *Nat. Protoc.* **2018**, *13*, 478 – 494.

[28] S. Kalkhof, A. Sinz, *Anal. Bioanal. Chem.* **2008**, *392*, 305 – 312.

[29] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Pérez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, Ş. Yılmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarnuczak, T. Ternent, A. Brazma, J. A. Vizcaíno, *Nucleic Acids Res.* **2019**, *47*, D442 – D450.