Chemical Science

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2019, **10**, 6443

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 21st February 2019
Accepted 23rd May 2019

DOI: 10.1039/c9sc00893d

rsc.li/chemical-science

A novel mass spectrometry-cleavable, phosphate-based enrichable and multi-targeting protein cross-linker†

Rong Huang,‡[ab] Wei Zhu,‡[a] Yue Wu,‡[a] Jiakang Chen,[a] Jianghui Yu,[ab] Biao Jiang,*[a] Hongli Chen *[a] and Wenzhang Chen*[a]

Chemical cross-linking mass spectrometry (XL-MS) is a powerful technology for obtaining protein structural information and studying protein–protein interactions. We report phospho-bisvinylsulfone (**pBVS**) as a novel water-soluble, MS-cleavable, phosphate-based enrichable and multi-targeting cross-linker. In this approach, the fragmentation of **pBVS** cross-linked peptides occurs *in situ* through retro-Michael addition. The phosphate group is successfully used as a small affinity tag to isolate cross-linked peptides from the highly abundant non-cross-linked peptides. In addition, the linker targets multiple types of amino acid residues, including cysteine, lysine and histidine. This method was applied to cross-link bovine serum albumin (BSA), myoglobin and Lbcpf1 demonstrating the ability to yield accurate and abundant information to facilitate protein structure elucidation.

Introduction

Chemical cross-linking-mass spectrometry (XL-MS) is emerging as a powerful tool to study protein–protein interactions at the proteome level and derive tertiary structural information about proteins and protein complexes.[1,2] With the increasing contributions of this technology to structural biology,[3–5] optimized XL-MS methods have been constantly developed and refined.[6–8] However, some technical challenges remain. The first challenge is the interpretation of cross-linking spectra on a large scale due to the exponential growth of the search space and poor spectral quality.[9–11] To address this limitation, MS-cleavable cross-linking reagents such as PIR,[12] DSSO,[13] DSBU, BuUrBu[6,14–16] and DAU[17] have been developed. A short, MS-cleavable linker, CDI, has also been reported recently.[18] The methods above can simplify the interpretation workflow and convert the computational complexity from O (n_2) to O (n), consequently generating far fewer false positive hits. The second challenge is that most protein–protein interactions are weak or transient, and non-cross-linked peptides are present in a large excess compared to cross-linked peptides. Due to the stochastic nature of MS experiments, most cross-linked peptides are not observable by MS.[10,19] To remove these non-cross-linked peptides, an affinity tag was incorporated into the cross-linking reagent to selectively enrich cross-linked peptide pairs.[20,21] For example, cross-linking reagents containing biotin as an affinity handle have been developed to boost the detection of cross-linked peptide pairs.[21–24] Another issue is that most cross-linking reagents target only one certain type of residue. For example, *N*-hydroxysuccinimide (NHS) esters, which are one of the most widely used cross-linking reagents, mainly react with lysine.[1,2,25] The specificity leads to fewer identifiable cross-linked peptide pairs, especially when there is no lysine available for cross-linking. Methods targeting acidic residues (Glu and Asp) have also been developed.[26,27] A SuFEx chemical cross-linker that can target different types of amino acid residues was recently reported.[28] It is necessary to develop an MS-cleavable, enrichable and multi-targeting cross-linker to study protein–protein interactions and elucidate protein structures.

Vinyl sulfones can react with several nucleophilic residues including cysteine, lysine and histidine.[29,30] Vinyl sulfone-based Michael addition is an attractive methodology for protein modification,[30,31] and we have paid much attention to the development of vinyl sulfones for conjugation with peptides and proteins.[32–35] Additionally, it is well-known that a high energy will induce retro-Michael addition, which triggers the preferential cleavage of the labile conjugation bond.[36,37] In this study, we intend to employ bisvinylsulfones as a novel class of MS-cleavable cross-linkers. In addition, although biotin is widely used as an affinity tag to enrich cross-linked peptides, the bulky biotin tag will certainly hinder the cross-linking efficiency due to steric effects and therefore result in fewer identifiable cross-linked peptide pairs.[38,39] Phosphate-based

[a]*Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, 393 Middle Huaxia Road, Pudong, Shanghai, 201210, China. E-mail: chenhl@shanghaitech.edu.cn; jiangbiao@shanghaitech.edu.cn; chenwzh1@shanghaitech.edu.cn*

[b]*University of Chinese Academy of Sciences, 19A Yuquan Road, Shijingshan District, Beijing, 100049, China*

† Electronic supplementary information (ESI) available: Experimental procedures and full spectroscopic data are available. See DOI: 10.1039/c9sc00893d

‡ These authors contributed equally to this work.

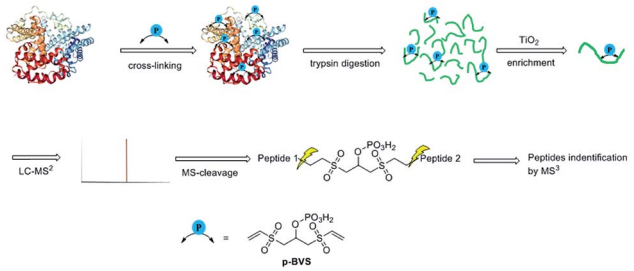


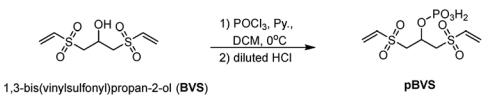
Fig. 1 XL-MS analysis workflow for identifying **pBVS** cross-linked peptides.

purification has been widely employed in proteomics analysis due to the affinity of phosphate to TiO_2.[40] In this study, a phosphate group was introduced into bisvinylsulfones in an attempt to selectively enrich cross-linked peptides using the TiO_2 method, and we expect the smaller affinity tag to minimize the steric impact of the cross-linker. Therefore, phosphobisvinylsulfone (**pBVS**) was developed as a novel class of MS-cleavable and enrichable cross-linkers (Fig. 1). To the best of our knowledge, this is the first use of a small phosphate group as an affinity tag. Cleaving the cross-linked peptides into two separate peptide chains through retro-Michael addition is also a novel distinct fragmentation mechanism. In addition, the linker demonstrates the ability to react with multiple types of amino acids under physiological conditions.

Results and discussion

pBVS was obtained from commercially available 1,3-bis(vinylsulfonyl)propan-2-ol (**BVS**) by one step phosphorylation (Scheme 1). Bovine serum albumin (BSA) was utilized to evaluate the cross-linking efficiency. The preliminary results showed that **pBVS** was less reactive to amino residues than NHS esters. However, **pBVS** shows good stability in aqueous solution, whereas NHS is sensitive to hydrolysis. After cross-linking with **pBVS**, the protein was digested,[41] and the resulting cross-linked peptides were enriched with TiO_2 (ref. 38) and analysed using collision-induced dissociation (CID)-MS^2-MS^3. 29 cross-linked peptides were identified, when the reaction was performed for 4 hours at room temperature and pH 7.4. To improve the efficiency of cross-linking, the reaction conditions were optimized with respect to time, temperature and pH value (ESI, Table S1†). 49 cross-linkers were found when the reaction was performed for 15 hours at 37 °C and pH 8.5.

As initially expected, the fragmentation of **pBVS** cross-linked peptides occurred *in situ* through retro-Michael addition under low-energy CID conditions. The cross-linked peptide P1–P2 was first fragmented into two pairs: $\alpha/\beta_{linker}/\beta_{linker\text{-}phos}$ and $\beta/\alpha_{linker}/$



Scheme 1 Synthesis of **pBVS**.

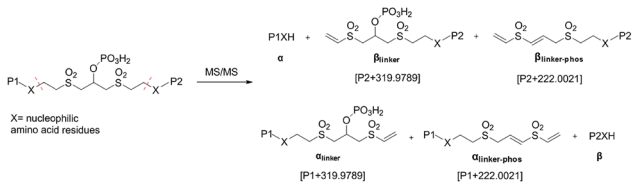


Fig. 2 Fragmentation products of **pBVS**-crosslinked peptides.

$\alpha_{linker\text{-}phos}$ (Fig. 2). A mass difference of 319.9789 was expected to be detected under low-energy MS/MS conditions. The α_{linker} and β_{linker} ions underwent further cleavage, which generated another mass tag (222.0021). These two mass tags (319.9789 and 222.0021) could be employed to trigger MS^3 analysis for the identification of cross-linked peptide sequences.

Fig. 3a and b display zoomed MS^2 spectra of a cross-linked α–β peptide ion (m/z = 1254.2218, z = 3) in the m/z ranges 940–1120 and 750–940, respectively, in which the two fragment ion pairs $\alpha_{linker}/\alpha_{linker\text{-}phos}/\alpha$ and $\beta_{linker}/\beta_{linker\text{-}phos}/\beta$ (z = 2) were detected. The mass difference between the ion α_{linker} and the ion α is 319.9744, equivalent to the exact mass of one **pBVS** linker with a mass error of approximately 0.005 Da. The mass difference between $\alpha_{linker\text{-}phos}$ and α is 221.9992, which is equivalent to the exact mass of one **pBVS** linker-phos with a mass error of approximately 0.003 Da. The α_{linker} and β ions were further fragmented by higher-energy collisional dissociation (HCD) (Fig. 3c and d). Subsequent database searching identified peptide α as NECFLSHKDDSPDLPK and β as LCVLHEKTPVSEK. In addition, the mass sum of the α_{linker} ion (1111.424, z = 2) and the β ion (770.4108, z = 2) is 3759.6384 Da, equivalent to the mass of the precursor ion 1254.2218 (z = 3, mass = 3759.642 Da) with an error less than 1 ppm. This finding suggests that the two fragment ions (ion 1111.424 and ion 770.4108) were formed by the cleavage of one labile bond of the precursor ion 1254.2218. We hence concluded that these two pairs of peptides were cross-linked by **pBVS**. The assignment of the ions in Fig. 3a and b and the detailed interpretation of α/

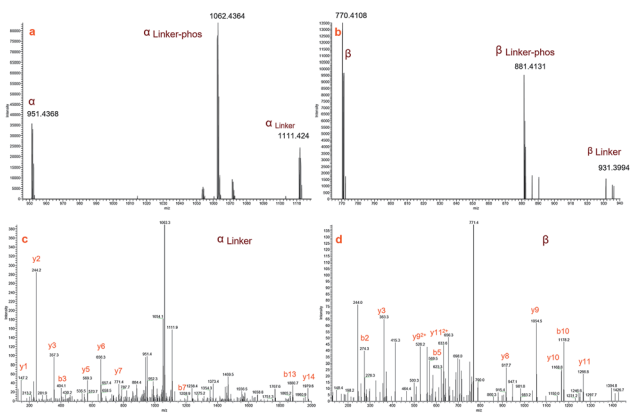
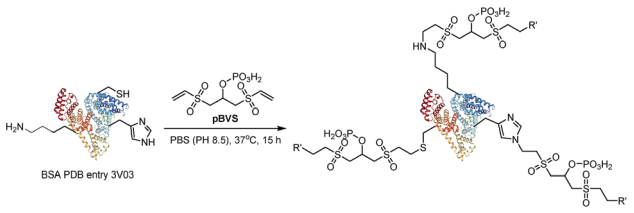


Fig. 3 MS^2 spectra of a representative cross-linked peptide pair (α–β, NECFLSHKDDSPDLPK–LCVLHEKTPVSEK): (a) MS^2 spectrum of α peptide ions; (b) MS^2 spectrum of β peptide ions; (c) MS^3 spectrum of α_{linker}; and (d) MS^3 spectrum of β.



Scheme 2 The reaction between **pBVS** with nucleophilic amino acid residues *via* Micheal addition.

$\alpha_{linker\text{-}phos}/\alpha_{linker}$ and $\beta/\beta_{linker\text{-}phos}/\beta_{linker}$ ions are shown in the ESI† (Tables S2–S8, Fig. S1–S6†).

The reactivity of **pBVS** towards different types of residues was also explored. As expected, **pBVS** is reactive to nucleophilic amino acid residues, including cysteine, lysine and histidine, *via* Michael addition (Scheme 2). As shown in Fig. 4a, the peptide YLQQCPFDEHVK contains one linker, and the existence of $b_5/b_6/y_7/y_8$ ions revealed that cross-linking occurred at C5 (Table S9†). The absence of [b_5-**pBVS**] ions and [y_{12}-**pBVS**] ions suggests that the **pBVS** linker is not dissociated from the peptide under high-energy collision dissociation (HCD) conditions. For the fragmentation of $\alpha_{linker}/\beta_{linker}$ ions, cleavages occurred along the peptide backbone rather than at the position between the **pBVS** linker and the peptide, and the feature can be utilized to localize the exact linked sites. Fig. 4b and c display the spectra of ions that were identified as LFTFHADICTLPDTEK and FWGKYLYEIAR, and the observation of a series of MS^3 ions revealed that cross-linking occurred at H5 and K4, respectively (Tables S10 and S11†).

Due to the linker's ability to facilitate both MS spectrum interpretation and sample purification, the identification of cross-linked pairs has been greatly improved compared to the results of previously reported methods.[18,28] A total of 49 unique cross-linked peptides, including 27 Lys–Lys, 20 Lys–His and 2 His–His, were detected by LC/MS/MS analysis from BSA

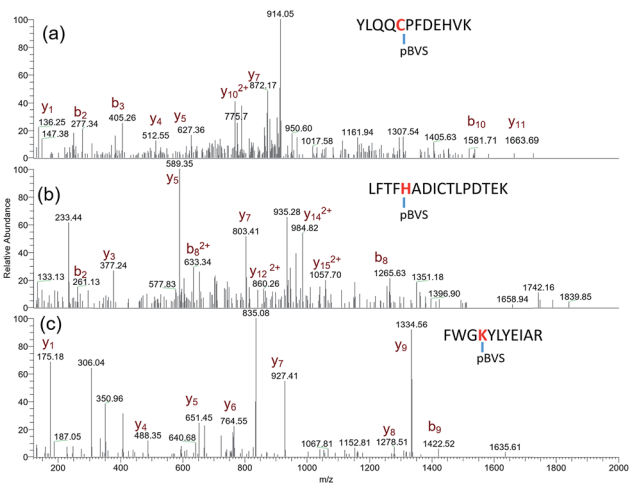


Fig. 4 **pBVS** as a multi-targeting cross linker: MS^3 spectra of YLQQCPFDEHVK + **pBVS** (a), LFTFHADICTLPDTEK + **pBVS** (b), and FWGKYLYEIAR + **pBVS** (c).

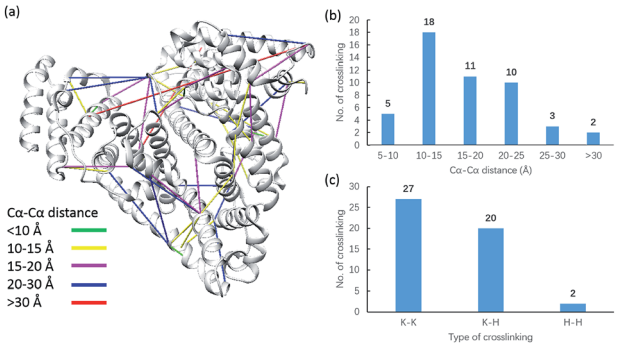


Fig. 5 Assessment of cross linkages with the BSA 3D structure: (a) cross-links mapped onto the crystal structures of BSA (PDB entry 3V03); (b) distribution of the Cα–Cα distance of the identified linkages; and (c) types of cross-linking sites identified from the **pBVS** cross-linked BSA sample, with numbers indicated on top.

(Fig. 5a–c, Table S12†). Unlike existing NHS cross-linkers that target only lysine, our method is able to target histidine in addition to lysine and thus demonstrates the ability to provide more information for protein structure elucidation. We also identified cross-linked peptides without enrichment to demonstrate the effects of the phosphate-based enrichment strategy. Twelve unique cross-linked peptides were detected (Table S13†), which confirmed the usefulness of the new enrichable strategy. To verify our data, the identified cross-linkages were mapped onto the crystal structures of BSA (Protein Data Bank entry: 3V03) to determine the Cα–Cα distance of the linked residues. Considering the spacer length of **pBVS**, which is ~15 Å, and the distances contributed by amino acid side chains, as well as the structural flexibility of the protein, the distance cut off for two connectable residues is set to 30 Å. Forty seven out of the 49 identified linkages (96%) were calculated to have Cα–Cα distances of less than 30 Å (Fig. 5). The two outliers could be explained by protein flexibility or dimerization. This result is consistent with the structural features of BSA, which indicates that the BSA structure was not disturbed by cross-linking. We also employed **pBVS** to investigate the

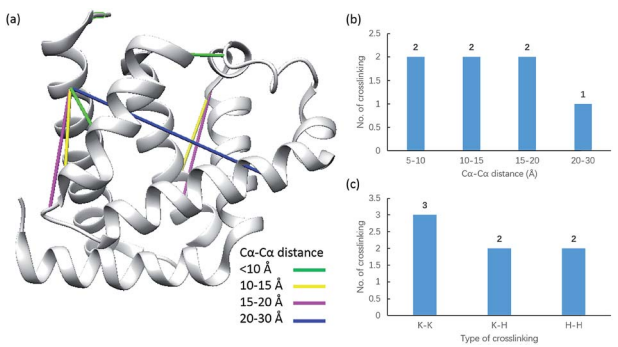


Fig. 6 Assessment of cross linkages with the myoglobin 3D structure: (a) cross-links mapped onto the crystal structures of myoglobin (PDB entry 1dwr); (b) distribution of the Cα–Cα distance of the identified linkages; and (c) types of cross-linking sites identified from the **pBVS** cross-linked myoglobin sample, with numbers indicated on top.

structure of myoglobin. Seven cross-linked peptides, including 3 Lys–Lys, 2 Lys–His and 2 His–His, were detected by LC/MS/MS analysis (Fig. 6a–c, Table S14†).

The distance data for all the linkages obtained from BSA and myoglobin except the outlier (60.5 Å) are plotted in Fig. 7a. The median value of the dataset is 15.2 Å, which is equal to the pBVS spacer length (~15 Å). The QQ plot in Fig. 7b shows that the dataset follows a normal distribution. This suggests that the distance error is mainly due to the flexibility of the amino acid side chains. One shortcoming of XL-MS for protein structure analysis is that the distance constraint is too loose to provide accurate structural information. The results of this study show that the method generates accurate distance information.

We further apply the reagent for structural elucidation of the Cpf1 protein from Lachnospiraceae bacteria (Lbcpf1), which is a single RNA-guided endonuclease, classified as a type V CRISPR system.[42,43] 19 unique linkages, including 13 Lys–Lys and 6 Lys–His, were detected (Fig. 8, Table S15†). When the cross-links were mapped onto the crystal structure (PDB entry: 5ID6), we found that 13 out of 19 linkages appeared in helical I (H1) and helical II (H2) domains, and 6 linkages were produced between H1 and its neighbouring OBD domain.[43]

To evaluate whether **pVBS** could be used in complex samples, we also applied the reagent to HeLa whole cell lysate. Unfortunately, the preliminary results show that most of the peptides detected were phosphopeptides rather than cross-linked peptides. This suggests that the method is more suitable to study samples in which protein phosphorylation is absent or less abundant.

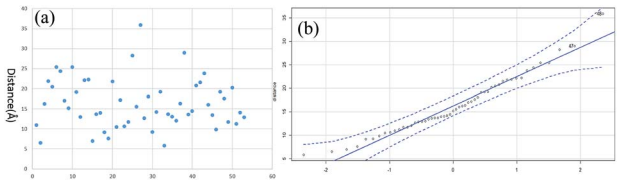


Fig. 7 The distribution of all the linkages: (a) Cα–Cα distance and (b) QQ plot.

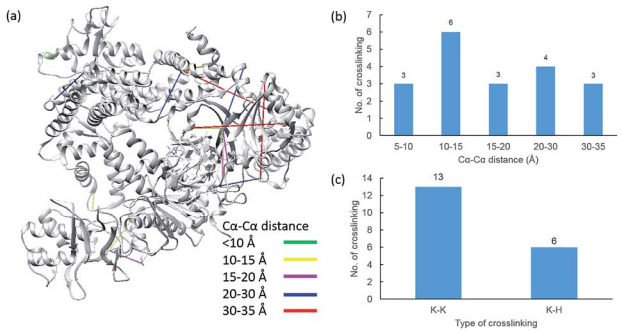


Fig. 8 Assessment of cross linkages with the Cpf1 3D structure: (a) cross-links mapped onto the crystal structures of Cpf1 (PDB entry 5ID6); (b) distribution of the Cα–Cα distance of the identified linkages; and (c) types of cross-linking sites identified from the **pBVS** cross-linked Cpf1 sample, with numbers indicated on top.

Conclusions

In this study, a novel MS-cleavable and enrichable cross-linking reagent, phospho-bisvinylsulfone (**pBVS**), was developed. This stable and water-soluble linker possesses an innovative modular scaffold and targets multiple types of amino acids (*i.e.*, lysine, histidine and cysteine) *via* Michael addition without any by-product. Under low-CID MS/MS conditions, cross-linked peptides are cleaved into two individual peptide chains through retro-Michael addition, which is a novel fragmentation pattern in XL-MS. A phosphate group as a small affinity tag is successfully used to isolate the cross-linked peptides from highly abundant non-cross-linked peptides. This novel strategy was applied to the cross-linking of the protein BSA, myoglobin and Lbcpf1, and the results demonstrated that **pBVS**-based linkages provide accurate and abundant information to facilitate protein structure elucidation. We also found one limitation of the method is that phosphopeptides will also be enriched and **pBVS** will not be suitable for complex samples with abundant phosphopeptides. Future studies will explore other types of vinyl sulfone-based cross-linkers suitable for complex samples.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Dr Heqiao Zhang for his help in data analysis of BSA cross-links. We also thank Dr Sheng Li, Shangshang Wang, Dr Lixia Zhao and Dr Yan Nie (ShanghaiTech University, SIAIS Protein & Gene Platform) for providing Lbcpf1.

Notes and references

1 C. Yu and L. Huang, *Anal. Chem.*, 2018, **90**, 144–165.
2 A. Sinz, *Angew. Chem., Int. Ed.*, 2018, **57**, 6390–6396.
3 P. Limpikirati, T. Liu and R. W. Vachet, *Methods*, 2018, **144**, 79–93.
4 C. Iacobucci, M. Gotze, C. H. Ihling, C. Piotrowski, C. Arlt, M. Schafer, C. Hage, R. Schmidt and A. Sinz, *Nat. Protoc.*, 2018, **13**, 2864–2889.
5 A. Sinz, *Anal. Bioanal. Chem.*, 2017, **409**, 33–44.
6 D. Pan, A. Brockmeyer, F. Mueller, A. Musacchio and T. Bange, *Anal. Chem.*, 2018, **90**, 10990–10999.
7 J. E. Horne, M. Walko, A. N. Calabrese, M. A. Levenstein, D. J. Brockwell, N. Kapur, A. J. Wilson and S. E. Radford, *Angew. Chem., Int. Ed.*, 2018, **57**, 16688–16692.
8 C. B. Gutierrez, S. A. Block, C. Yu, S. M. Soohoo, A. S. Huszagh, S. D. Rychnovsky and L. Huang, *Anal. Chem.*, 2018, **90**, 7600–7607.
9 F. Yu, N. Li and W. Yu, *J. Proteome Res.*, 2017, **16**, 3942–3952.
10 F. Liu, D. T. Rijkers, H. Post and A. J. Heck, *Nat. Methods*, 2015, **12**, 1179–1184.

11 L. Lu, R. J. Millikin, S. K. Solntsev, Z. Rolfs, M. Scalf, M. R. Shortreed and L. M. Smith, *J. Proteome Res.*, 2018, **17**, 2370–2376.
12 X. Tang, G. R. Munske, W. F. Siems and J. E. Bruce, *Anal. Chem.*, 2005, **77**, 311–318.
13 A. Kao, C. L. Chiu, D. Vellucci, Y. Yang, V. R. Patel, S. Guan, A. Randall, P. Baldi, S. D. Rychnovsky and L. Huang, *Mol. Cell. Proteomics*, 2011, **10**, M110.002212.
14 D. L. Smith, M. Gotze, T. K. Bartolec, G. Hart-Smith and M. R. Wilkins, *Anal. Chem.*, 2018, **90**, 9101–9108.
15 C. Arlt, M. Gotze, C. H. Ihling, C. Hage, M. Schafer and A. Sinz, *Anal. Chem.*, 2016, **88**, 7930–7937.
16 M. Q. Muller, F. Dreiocker, C. H. Ihling, M. Schafer and A. Sinz, *Anal. Chem.*, 2010, **82**, 6958–6968.
17 C. Iacobucci, C. Piotrowski, A. Rehkamp, C. H. Ihling and A. Sinz, *J. Am. Soc. Mass Spectrom.*, 2019, **30**, 139–148.
18 C. Hage, C. Iacobucci, A. Rehkamp, C. Arlt and A. Sinz, *Angew. Chem., Int. Ed.*, 2017, **56**, 14551–14555.
19 M. R. Hoopmann, A. Zelter, R. S. Johnson, M. Riffle, M. J. MacCoss, T. N. Davis and R. L. Moritz, *J. Proteome Res.*, 2015, **14**, 2190–2198.
20 X. Zhong, A. T. Navare, J. D. Chavez, J. K. Eng, D. K. Schweppe and J. E. Bruce, *J. Proteome Res.*, 2017, **16**, 720–727.
21 X. Tang and J. E. Bruce, *Mol. BioSyst.*, 2010, **6**, 939–947.
22 A. N. Holding, *Methods*, 2015, **89**, 54–63.
23 R. M. Kaake, X. Wang, A. Burke, C. Yu, W. Kandur, Y. Yang, E. J. Novtisky, T. Second, J. Duan, A. Kao, S. Guan, D. Vellucci, S. D. Rychnovsky and L. Huang, *Mol. Cell. Proteomics*, 2014, **13**, 3533–3543.
24 S. C. Alley, F. T. Ishmael, A. D. Jones and S. J. Benkovic, *J. Am. Chem. Soc.*, 2000, **122**, 6126–6127.
25 J. Rappsilber, *Eur. J. Org. Chem.*, 2011, **173**, 530–540.
26 C. B. Gutierrez, C. Yu, E. J. Novitsky, A. S. Huszagh, S. D. Rychnovsky and L. Huang, *Anal. Chem.*, 2016, **88**, 8315–8322.
27 A. Leitner, L. A. Joachimiak, P. Unverdorben, T. Walzthoeni, J. Frydman, F. Forster and R. Aebersold, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 9455–9460.
28 B. Yang, H. Wu, P. D. Schnier, Y. Liu, J. Liu, N. Wang, W. F. DeGrado and L. Wang, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 11162–11167.
29 O. Koniev and A. Wagner, *Chem. Soc. Rev.*, 2015, **44**, 5495–5551.
30 J. Morales-Sanfrutos, J. Lopez-Jaramillo, M. Ortega-Munoz, A. Megia-Fernandez, F. Perez-Balderas, F. Hernandez-Mateo and F. Santoyo-Gonzalez, *Org. Biomol. Chem.*, 2010, **8**, 667–675.
31 A. Maruani, D. A. Richards and V. Chudasama, *Org. Biomol. Chem.*, 2016, **14**, 6165–6178.
32 H. Chen, R. Huang, Z. Li, W. Zhu, J. Chen, Y. Zhan and B. Jiang, *Org. Biomol. Chem.*, 2017, **15**, 7339–7345.
33 R. Huang, Z. Li, P. Ren, W. Chen, Y. Kuang, J. Chen, Y. Zhan, H. Chen and B. Jiang, *J. Struct. Biol.*, 2018, **2018**, 829–836.
34 R. Huang, Z. Li, Y. Sheng, J. Yu, Y. Wu, Y. Zhan, H. Chen and B. Jiang, *Org. Lett.*, 2018, **20**, 6526–6529.
35 Z. Li, R. Huang, H. Xu, J. Chen, Y. Zhan, X. Zhou, H. Chen and B. Jiang, *Org. Lett.*, 2017, **19**, 4972–4975.
36 T. Selvi and S. Velmathi, *J. Org. Chem.*, 2018, **83**, 4087–4091.
37 R. Kobetic, S. Kazazic, B. Kovacevic, Z. Glasovac, L. Krstulovic, M. Bajic and B. Zinic, *J. Am. Soc. Mass Spectrom.*, 2015, **26**, 833–842.
38 E. V. Petrotchenko and C. H. Borchers, *Mass Spectrom. Rev.*, 2010, **29**, 862–876.
39 S. M. Chowdhury, X. X. Du, N. Tolic, S. Wu, R. J. Moore, M. U. Mayer, R. D. Smith and J. N. Adkins, *Anal. Chem.*, 2009, **81**, 5524–5532.
40 L. R. Yu, Z. Zhu, K. C. Chan, H. J. Issaq, D. S. Dimitrov and T. D. Veenstra, *J. Proteome Res.*, 2007, **6**, 4150–4162.
41 J. R. Wisniewski, A. Zougman, N. Nagaraj and M. Mann, *Nat. Methods*, 2009, **6**, 359–U360.
42 B. Zetsche, J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin and F. Zhang, *Cell*, 2015, **163**, 759–771.
43 D. Dong, K. Ren, X. Qiu, J. Zheng, M. Guo, X. Guan, H. Liu, N. Li, B. Zhang, D. Yang, C. Ma, S. Wang, D. Wu, Y. Ma, S. Fan, J. Wang, N. Gao and Z. Huang, *Nature*, 2016, **532**, 522–526.