Chemical Modification of Peptides and Proteins Using Spirooxindole Oxirane Derivatives

doi.org/10.1002/adsc.202300578

Advanced Synthesis & Catalysis

Chemical Modification of Peptides and Proteins Using Spirooxindole Oxirane Derivatives

Santosh S. Chavan,[a] Hidetoshi Saze,[b] and Fujie Tanaka[a,*]

[a] Chemistry and Chemical Bioengineering Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Okinawa 904-0495, Japan
phone: +81-98-966-1552; fax: +81-98-966-1064
E-mail: ftanaka@oist.jp

[b] Plant Epigenetics Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Okinawa 904-0495, Japan

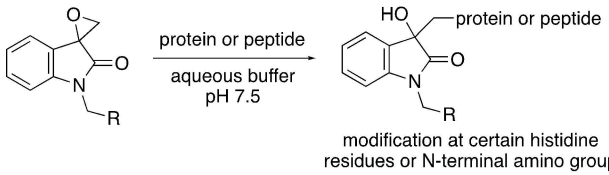
Manuscript received: May 31, 2023; Version of record online: June 29, 2023

Supporting information for this article is available on the WWW under https://doi.org/10.1002/adsc.202300578

Abstract: Methods for chemical modification of proteins and peptides at certain positions are required for the preparation of antibody-drug conjugates, protein- or peptide-biotin conjugates, and conjugates used as therapeutics or as tools in biomedical research. We have developed chemical modification methods that use spirooxindole oxirane derivatives for the synthesis of protein- and peptide-conjugates. In most cases, the epoxide group of the spirooxindole oxirane derivatives reacted to form a covalent bond with certain histidine residues and/or the N-terminal amino group of peptides and proteins. The modification reactions of various proteins resulted in the formation of proteins modified at only one or two positions.

Keywords: addition reaction; protein modifications; peptide modifications; spiro compounds

Chemical modification reactions of proteins and peptides are necessary for the synthesis of protein conjugates and peptide conjugates that are used as therapeutics, diagnostics, and tools for biomedical research.[1] Traditional protein modification reactions include, for example, reactions at lysine amino groups with N-hydroxysuccinimide (NHS) ester derivatives.[1a] Reactions with NHS derivatives often result in the formation of a mixture of products modified at different lysine residues. The mixture of the species can cause difficulties in characterization and lack of reproducibility in quantitative analyses. It is especially critical that the modified proteins used as therapeutics or in diagnostics can be purified as a single species. Thus, in recent years, there has been interest in the development of methods for selective modification at a single position or only a few positions within a protein.[1] Selective modifications have been performed using ligand-assisted or linchpin-assisted (amino group-interaction assisted) strategies, in which only targeted protein and targeted positions are modified based on the design of the protein-ligand or protein-linchpin interaction.[1,2] Alternatively, selective modifications of the N-terminal amino group, which is a single unique site within a polypeptide chain of a protein, and/or of amino acid residues that are present at only one or a few sites on the surface of a protein have also resulted in modifications at a single position or only a few positions within a protein.[1a,f,n,s] Here we report a chemical modification reaction strategy that uses spirooxindole oxirane derivatives to modify N-terminal amino groups and/or certain histidines under mild conditions (Scheme 1).



Scheme 1. Chemical modifications of peptides and proteins with spirooxindole oxirane derivatives as modification molecules.

In synthetic chemistry, epoxide derivatives have been used for various reactions including reactions with nucleophiles.[3] In protein modification reactions, alkyl-substituted epoxide derivatives that are linked with ligands of the target proteins or moieties that function as linchpins have been used for modification of certain histidine residues of the target proteins.[2] Alkyl-substituted epoxides have also been used to identify proteins that have exceptionally highly nucleophilic residues.[4] These previous reports indicate that alkyl-substituted epoxides do not react with most proteins under aqueous buffered conditions unless strategies are used that locate the epoxide in proximity to the reacting residue on the protein. In fact, when reactions of 1,2-epoxyhexane with proteins in aqueous solutions were tested, proteins were not modified or the modified products were formed in less than 3% yield. Thus, we sought designs of epoxide-derived modification molecules that modify various peptides and proteins at only a single or a few certain positions within a peptide or protein.

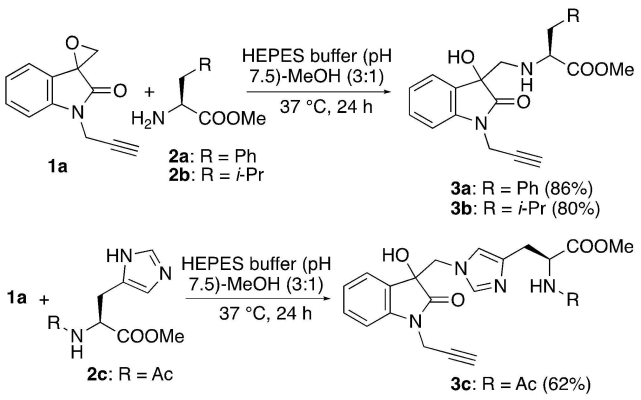
We hypothesized that spirooxindole oxirane derivatives would serve as molecules used for modifications of proteins and peptides. We reasoned that spirooxindole oxirane derivatives would react at histidine residues, N-terminal amino groups, and/or other nucleophilic residues depending on the structure of peptide or protein, the structure of the spirooxindole oxirane derivative (i.e., substituents and functional moieties linked with the derivative), and reaction conditions. There are fewer histidine residues on protein surfaces than lysine residues and there is only a single N-terminal amino group per polypeptide chain. Thus, reactions at histidine residues and/or at the N-terminal amino group should result in more selective modification than reactions with lysine residues.[1g,h] However, the development of reactions at certain histidine residues with modification molecules without the need for protein-ligand or protein-linchpin interactions and without the need of additional reagents or catalysts has lagged behind,[1b,g] and as such further development is required.

In our design, the epoxide moiety is conjugated with a lactam amide carbonyl group, which is electron-withdrawing, and thus the reactions of proteins and peptides with the epoxide should proceed efficiently. Functional groups introduced to proteins and peptides would be linked to the nitrogen of the lactam amide of the oxindole structure. Tuning of the epoxide reactivity should also be possible by the design of the linker structure that connects the oxindole with the functional moieties to be introduced into proteins or peptides. Tuning of the reactivities and/or the steric bulk of the spirooxindole oxirane derivatives should allow selective modifications at certain histidines or other positions in a protein without reacting at functionally important histidine residues.

To evaluate the use of spirooxindole oxirane derivatives in modifications of proteins and peptides, first, reactions of **1a** with amino acid derivatives, phenylalanine methyl ester (**2a**), leucine methyl ester (**2b**), and N-acetyl histidine methyl ester (**2c**) were tested in aqueous buffer (pH 7.5)-MeOH (3:1) (Scheme 2). In the reactions with **2a** and with **2b**, products **3a** and **3b** were formed, respectively, in which the amino group of the amino acid ester reacted with the less substituted side of the epoxide moiety. In the reaction with **2c**, product **3c** was obtained. In this case, the imidazole group of the histidine derivative reacted with the epoxide moiety. These results suggest that the spirooxindole oxirane derivatives will react with proteins and peptides to afford conjugates.

In previously reported reactions of spirooxindole oxirane derivatives with various nucleophiles, the bond formation usually occurs at the benzylic position of the indole structure.[5] There are only a few reports of reactions at the methylene side of the epoxide ring of spirooxindole oxirane derivatives.[6] Our data showed that the reactions of **1a** with amino acid derivatives in aqueous buffered conditions resulted in the N–C bond formation at the methylene side of the epoxide.

Next, reactions of **1a** with short peptides were examined (Table 1 and Figure 1; see also Supporting Information for additional conditions). When peptide AIKVF-NH_2 was used in the reaction with **1a** under aqueous conditions (pH 7.0–8.0), the N-terminal amino group was reacted with **1a**; bond formation between the lysine side chain amino group and **1a** was not detected in the MS/MS analyses (Table 1, entries 1–3). The reaction of AIKVF-NH_2 with **1a** in various buffers containing Tris, HEPES, PIPES, TAPS, or sodium phosphate at pH 7.0–8.0 afforded the modification product (Supporting Information). Tris and TAPS have a primary amine group and a secondary amine group, respectively, and these amino groups did not react with **1a** or result in its decomposition. The reaction of peptide AIKVF-NH_2 with **1a** in 10% DMF/20 mM



Scheme 2. Reactions of amino acid esters with **1a**.

Table 1. Reactions of **1 a** with short peptides.[a]

entry	peptide	yield (%)	modified site
1	AIKVF-NH_2	mono 93	N-terminus
2[b]	AIKVF-NH_2	mono 65	N-terminus
3[c]	AIKVF-NH_2	mono 92	N-terminus
4	AIRVF-NH_2	mono 58	N-terminus
5	LAFKVPEGDF-NH_2	mono 22	N-terminus
6	AIHVF-NH_2	mono 85, di 1	N-terminus, His3
7	YGGFL-OH	mono 26	N-terminus

[a] Conditions: [**1 a**] 2.5 mM and [peptide] 50 μM in 10% DMF/20 mM Tris-HCl (pH 7.5) at 37 °C for 24 h. Yield and modification site were determined by mass analyses; mono = mono-modification product(s); di = di-modification product(s).
[b] At pH 7.0.
[c] At pH 8.0.

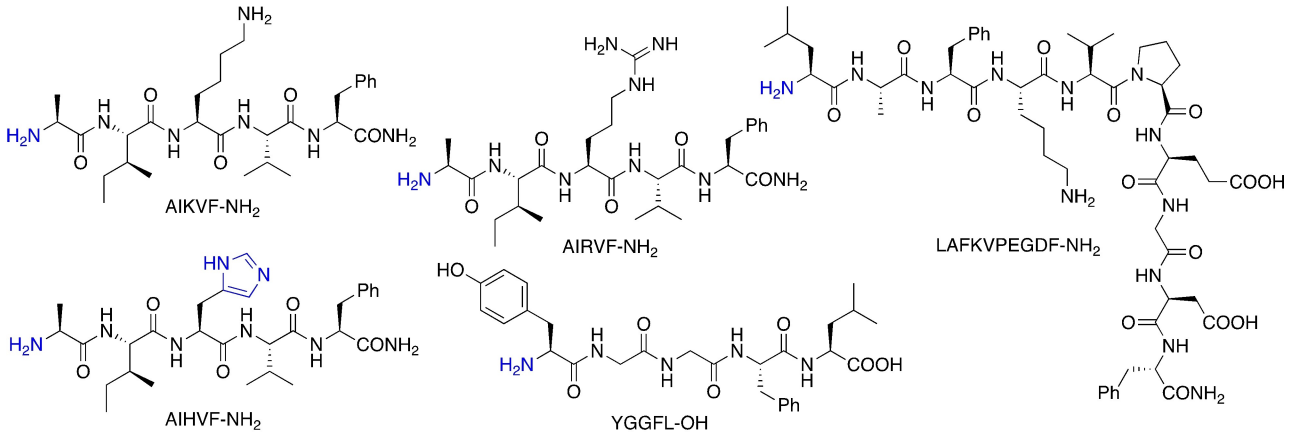


Figure 1. Peptides used for the reaction with **1 a**. Positions modified with **1 a** are indicated in blue. See Table 1.

Tris-HCl buffer (pH 7.5) at 37 °C for 24 h afforded the mono-modification product in 93% yield based on the mass analysis, and no di-modification product was observed (Table 1, entry 1). Reactions of other peptides were also performed under the same conditions (Table 1, entries 4–7). For the histidine-containing peptide, the bond formations were observed at histidine and at the N-terminus (Table 1, entry 6).

For the modifications of proteins with NHS ester derivatives, the use of 10 to 70 equivalents of the NHS esters relative to protein molecule has often been recommended.[7] The concentration of **1 a** at 2.5 mM used in the modification reactions is comparable to those typically used in modification reactions with NHS esters.

Next, biotin conjugates **1 b** and **1 c** were synthesized and were tested for modification reactions with various proteins in aqueous buffers (Table 2 and Figure 2). In many cases, the reactions resulted in histidine modification. Accessibility and/or reactivity (or pKa) of the histidine residues may determine whether or not reaction occurs. For example, α-lactalbumin has three histidine residues, and only two were modified (Table 2, entries 7 and 8). For cytochrome C and myoglobin, only certain histidine residues were modified, and the heme-liganded histidine residues were not modified (Table 2, entries 13–16). Depending on the protein, the N-terminus and/or lysine side chain amino groups were also modified. For example, in the reaction of α-chymotrypsin, only the N-terminal amino group of chain A was modified (Table 2, entries 17 and 18). The N-terminal amino groups of chains B and C were not modified, and neither were any histidine or lysine residues including the protease catalytic active site histidine.[8] Ubiquitin has lysine residues critical for ubiquitylation and polyubiquitylation.[9] In the reactions of ubiquitin, a single histidine was modified, and no modification at lysine residues was detected (Table 2, entries 3 and 4). Notably, in the reaction of β-lactoglobulin, the histidine residues were modified, but the modification of the thiol of a cysteine residue was not detected (Table 2, entries 9 and 10).

Modification reactions of the Fab of an anti-CD20 antibody[10] with **1 b** and with **1 c** were also examined (Table 3 and Figure 2k). The Fab has four histidine residues in the H chain and three histidine residues in the L chain. The results indicated that only certain

Table 2. Reactions of proteins with **1 b** and with **1 c**.[a]

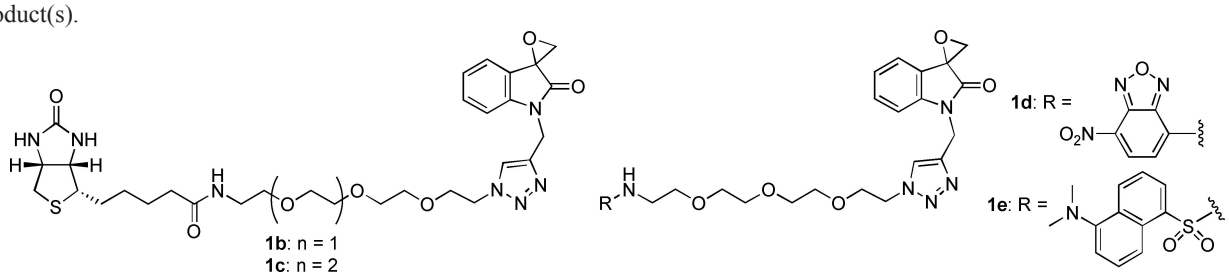
entry	protein	1	yield (%)	modified site
1	lysozyme	**1 b**	mono 23, di 5	His15
2	lysozyme	**1 c**	mono 24, di 4	His15, Lys33
3	ubiquitin	**1 b**	mono 50, di 3	His68
4	ubiquitin	**1 c**	mono 49, di 4	His68
5	insulin	**1 b**	mono 42, di 10	chain B His5, chain B His10
6	insulin	**1 c**	mono 43, di 11	chain B His5, chain B His10
7	α-lactalbumin	**1 b**	mono 29, di 8	His68, His107
8	α-lactalbumin	**1 c**	mono 30, di 9	His68, His107
9	β-lactoglobulin	**1 b**	mono 35, di 13	His146, His161
10	β-lactoglobulin	**1 c**	mono 40, di 22, tri 6	His146, His161, Lys135
11	ribonuclease A	**1 b**	mono 35, di 10	His48, His105
12	ribonuclease A	**1 c**	mono 39, di 10	His105, His119
13	cytochrome C	**1 b**	mono 19, di < 3	His33
14	cytochrome C	**1 c**	mono 24, di 4	His26
15	myoglobin	**1 b**	mono 40, di 11	His24, His36
16	myoglobin	**1 c**	mono 40, di 12	His24, His36
17	α-chymotrypsin	**1 b**	mono 57, di 0	chain A N-terminus
18	α-chymotrypsin	**1 c**	mono 50, di 0	chain A N-terminus
19	α-chymotrypsinogen A	**1 b**	mono 50, di 2	N-terminus
20	α-chymotrypsinogen A	**1 c**	mono 68, di 17	N-terminus, Lys93

[a] Conditions: [**1b** or **1c**] 2.5 mM and [protein] 50 μM in 5% DMF/20 mM Tris-HCl (pH 7.5) at 37 °C for 24 h. Yield and modification site were determined by mass analyses; mono = mono-modification product(s); di = di-modification product(s); tri = tri-modification products.

Table 3. Reactions of an antibody Fab with **1 b** and with **1 c**.[a]

entry	1 (conc)	yield (%)	modified site
1	**1 b** (1.5 mM)	mono 24, di < 3	H chain His35
2	**1 b** (2.5 mM)	mono 40, di 4	H chain His35, L chain His188
3	**1 c** (1.5 mM)	mono 19, di < 1	L chain His188
4	**1 c** (2.5 mM)	mono 42, di 8	H chain His35, L chain His188

[a] Conditions: [**1 b** or **1 c**] 1.5 mM or 2.5 mM as indicated and [Fab] 29 μM in 5% DMF/20 mM Tris-HCl (pH 7.5) at 37 °C for 24 h. Yield and modification site were determined by mass analyses; mono = mono-modification product(s); di = di-modification product(s).



histidine residues were modified. No lysine or N-terminal modification was detected.

Further, compounds **1 d** and **1 e**, which have fluorescent moieties, were synthesized and used for modification of α-chymotrypsinogen A and ubiquitin. SDS/PAGE analysis of the modified proteins showed that the nitrobenzoxadiazole and the dansyl groups respectively were introduced to the proteins and that the modified proteins were detected by the fluorescence of the fluorescent moieties (Supporting Information).

In summary, we developed a strategy for selective chemical modifications of proteins and peptides that uses spirooxindole oxirane derivatives. With these modification molecules, one or a few histidine residues or the N-terminal amino group of proteins and peptides were selectively modified in most cases under mild aqueous buffered conditions without the need of additional reagents or catalysts. We are working to expand

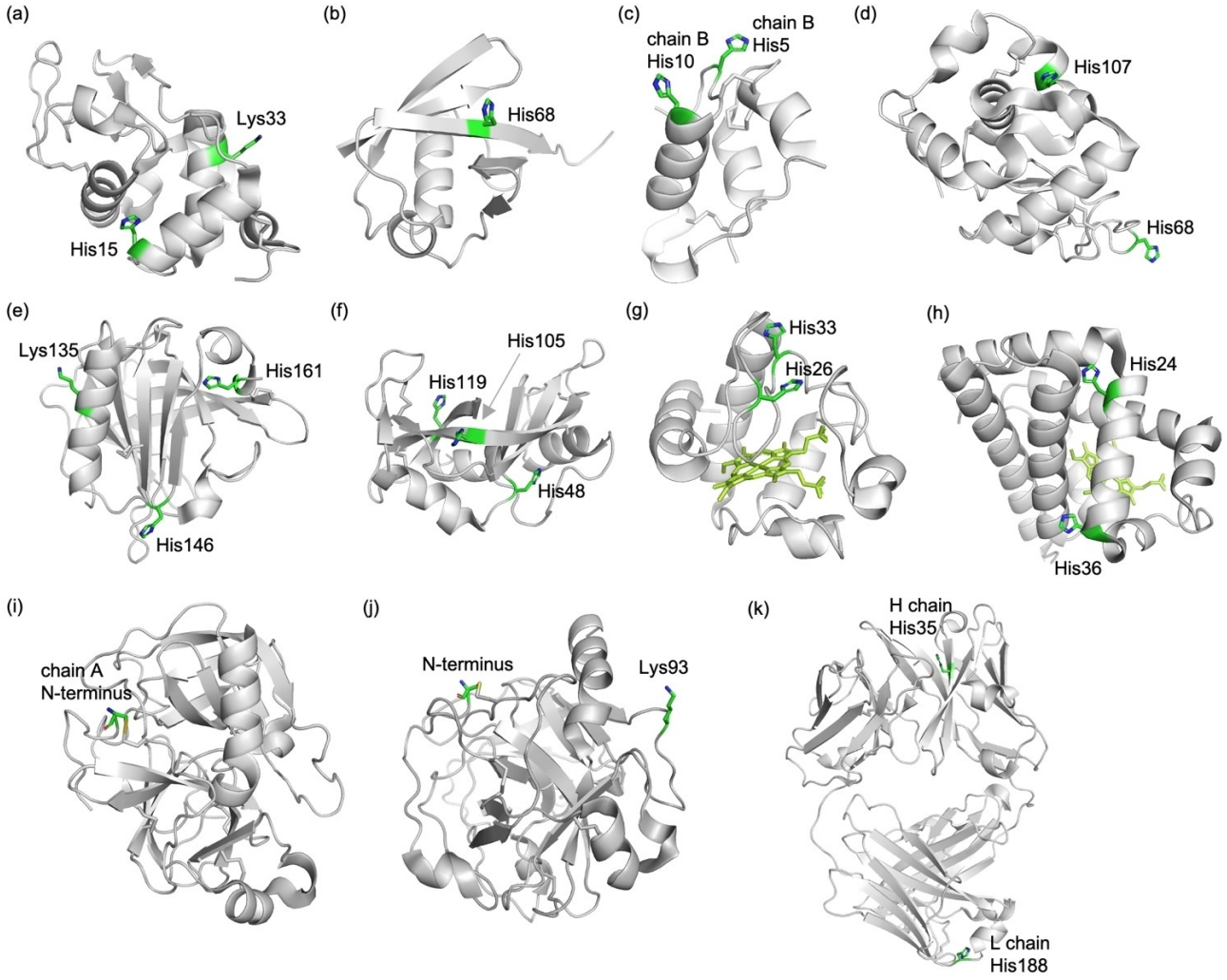


Figure 2. Positions modified with **1 b** and **1 c** are shown on previously reported PDB structures. See Tables 2 and 3. Proteins and PDB codes are the following: (a) lysozyme, 1dpx; (b) ubiquitin, 1ubq; (c) insulin, 6o17; (d) α-lactalbumin, 1f6r; (e) β-lactoglobulin, 3npo; (f) ribonuclease A, 1 fs3; (g) cytochrome C, 1hrc; (h) myoglobin, 1npg; (i) α-chymotrypsin, 1yph; (j) α-chymotrypsinogen A, 1ex3; and (k) anti-CD20 antibody Fab, 6vja.

the applications of the method and to improve the yields and selectivities, and the results will be reported in due course.

Experimental Section

General Procedure of Modification Reactions of Proteins with 1 b

A solution of protein (1.0 mM in 20 mM Tris HCl buffer, pH 7.5, freshly prepared, 5.0 μL) was added to 20 mM Tris HCl buffer, pH 7.5 (90.0 μL), followed by **1 b** (50 mM in DMF, 5.0 μL) at room temperature (25 °C), and the mixture was vortexed for 10 s. The mixture contained protein (50 μM) and **1 b** (2.5 mM) in 5% DMF/20 mM Tris HCl buffer, pH 7.5. The mixture was kept with shaking at 37 °C for 24 h. The mixture was applied to PD Spin Trap G-25 gel filtration column device (Cytiva 28918004), which was washed with 20 mM Tris HCl buffer, pH 7.5 before the use, and the protein was recovered as a solution in 20 mM Tris HCl buffer, pH 7.5 (100 μL), according to the instruction of the device provided by the maker. The solution (4.0 μL) was diluted with 0.1% formic acid in water (80 μL) and was used for the mass analysis to determine the yields of the modification products.

Acknowledgements

We thank Dr. Michael Chandro Roy, Research Support Division, Okinawa Institute of Science and Technology Graduate University (OIST) for mass analyses of modification molecules and for the synthesis of short peptides. We also thank Dr. Yoshitoshi Hirao and Dr. Alejandro Villar-Briones, Research Support Division, OIST for useful discussions and support for mass analyses of proteins and peptides, including proteomic analyses performed to analyze modification reactions. We also thank Sysmex Corporation for providing the antibody Fab protein.

This study was supported by the Okinawa Institute of Science and Technology Graduate University.

References

[1] a) O. Boutureira, G. J. L. Bernardes, *Chem. Rev.* **2015**, *115*, 2174; b) J. N. de Gruyter, L. R. Malins, P. S. Baran, *Biochemistry* **2017**, *56*, 3863; c) E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira, G. J. L. Bernardes, *Nat. Chem. Rev.* **2019**, *3*, 147; d) N. Krall, F. P. da Cruz, O. Boutureira, G. J. L. Bernardes, *Nat. Chem.* **2016**, *8*, 103; e) S. J. Walsh, J. D. Bargh, F. M. Dannheim, A. R. Hanby, H. Seki, A. J. Counsell, X. Ou, E. Fowler, N. Ashman, Y. Takada, A. Isidro-Llobet, J. S. Parker, J. S. Carroll, D. R. Spring, *Chem. Soc. Rev.* **2021**, *50*, 1305; f) C. B. Rosen, M. B. Francis, *Nat. Chem. Biol.* **2017**, *13*, 697; g) S. Jia, D. He, C. J. Chang, *J. Am. Chem. Soc.* **2019**, *141*, 7294; h) X. Chen, F. Ye, X. Luo, X. Liu, J. Zhao, S. Wang, Q. Zhou, G. Chen, P. Wang, *J. Am. Chem. Soc.* **2019**, *141*, 18230; i) J. C. Maza, D. L. V. Bader, L. Xiao, A. M. Marmelstein, D. D. Brauer, A. M. ElSohly, M. J. Smith, S. W. Krska, C. A. Parish, M. B. Francis, *J. Am. Chem. Soc.* **2019**, *141*, 3885; j) K. Maruyama, T. Ishiyama, Y. Seki, K. Sakai, T. Togo, K. Oisaki, M. Kanai, *J. Am. Chem. Soc.* **2021**, *143*, 19844; k) M. T. Taylor, J. E. Nelson, M. G. Suero, M. J. Gaunt, *Nature* **2018**, *562*, 563; l) A. H. Christian, S. Jia, W. Cao, P. Zhang, A. T. Meza, M. S. Sigman, C. J. Chang, F. D. Toste, *J. Am. Chem. Soc.* **2019**, *141*, 12657; m) J. C. Vantourout, S. Rao Adusumalli, K. W. Knouse, D. T. Flood, A. Ramirez, N. M. Padial, A. Istrate, K. Maziarz, J. N. deGruyter, R. R. Merchant, J. X. Qiao, M. A. Schmidt, M. J. Deery, M. D. Eastgate, P. E. Dawson, G. J. L. Bernardes, P. S. Baran, *J. Am. Chem. Soc.* **2020**, *142*, 17236; n) Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohma, K. Oisaki, M. Kanai, *J. Am. Chem. Soc.* **2016**, *138*, 10798; o) S. J. Tower, W. J. Hetcher, T. E. Myers, N. J. Kuehl, M. T. Taylor, *J. Am. Chem. Soc.* **2020**, *142*, 9112; p) C. R. Hoopes, F. J. Garcia, A. M. Sarkar, N. J. Kuehl, D. T. Barkan, N. L. Collins, G. E. Meister, T. R. Bramhall, C.-H. Hsu, M. D. Jones, M. Schirle, M. T. Taylor, *J. Am. Chem. Soc.* **2022**, *144*, 6227; q) B. Snella, B. Grain, J. Vicogne, F. Capet, B. Wiltschi, O. Melnyk, V. Agouridas, *Angew. Chem. Int. Ed.* **2022**, *61*, e202204992; r) T. S. Howard, R. D. Cohen, O. Nwajiobi, Z. P. Muneeswaran, Y. E. Sim, N. N. Lahankar, J. T.-H. Yeh, M. Raj, *Org. Lett.* **2018**, *20*, 5344; s) N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942.

[2] a) G. Chen, A. Heim, D. Riether, D. Yee, Y. Milgrom, M. A. Gawinowicz, D. Sames, *J. Am. Chem. Soc.* **2003**, *125*, 8130; b) Y. Takaoka, H. Tsutsumi, N. Kasagi, E. Nakata, I. Hamachi, *J. Am. Chem. Soc.* **2006**, *128*, 3273; c) S. R. Adusumalli, D. G. Rawale, U. Singh, P. Tripathi, R. Paul, N. Kalra, R. K. Mishra, S. Shukla, V. Rai, *J. Am. Chem. Soc.* **2018**, *140*, 15114.

[3] a) Y. Wang, J. Chai, C. You, J. Zhang, X. Mi, L. Zhang, S. Luo, *J. Am. Chem. Soc.* **2020**, *142*, 3184; b) A. K. Hubbell, A. M. LaPointe, J. R. Lamb, G. W. Coates, *J. Am. Chem. Soc.* **2019**, *141*, 2474; c) M. J. Micheltree, J. W. Stevenson, A. Pisipati, A. G. Myers, *J. Am. Chem. Soc.* **2021**, *143*, 6829; d) M. Magre, E. Paffenholz, B. Maity, L. Cavallo, M. Rueping, *J. Am. Chem. Soc.* **2020**, *142*, 14286.

[4] M. J. Evans, A. Saghatelian, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol.* **2005**, *23*, 1303.

[5] a) G. Zhu, G. Bao, Y. Li, W. Sun, J. Li, L. Hong, R. Wang, *Angew. Chem. Int. Ed.* **2017**, *56*, 5332; b) S. Hajra, S. Roy, *Org. Lett.* **2020**, *22*, 1458; c) S. Hajra, S. Roy, A. S. Mondal, *Adv. Synth. Catal.* **2020**, *362*, 5475; d) G. Zhu, Y. Li, G. Bao, W. Sun, L. Huang, L. Hong, R. Wang, *ACS Catal.* **2018**, *8*, 1810; e) L. Wu, Q. Shao, L. Kong, J. Chen, Q. Wei, W. Zhang, *Org. Chem. Front.* **2020**, *7*, 862.

[6] a) B. Zhang, Y. Li, G. Bao, G. Zhu, J. Li, J. Wang, B. Zhang, W. Sun, L. Hong, R. Wang, *Green Chem.* **2017**, *19*, 2107; b) F.-R. Zhang, N.-W. Wan, J.-M. Ma, B.-D. Cui, W.-Y. Han, Y.-Z. Chen, *ACS Catal.* **2021**, *11*, 9066.

[7] a) E. D. Bayer, M. Wilchek, Protein Biotinylation, Methods in Enzymology, **1990**, *184*, pp138–160; b) Instructions for the use of commercially available NHS ester derivatives (such as biotin-linked NHS derivatives) provided by the makers.

[8] C. Branden, J. Tooze, Introduction to Protein Structure 2nd Ed. Garland Publishing, Inc. **1999**, pp 210–212.

[9] C. Behrends, J. W. Harper, *Nat. Struct. Mol. Biol.* **2011**, *18*, 520.

[10] L. Rougé, N. Chiang, M. Steffek, C. Kugel, T. I. Croll, C. Tam, A. Estevez, C. P. Arthur, C. M. Koth, C. Ciferri, E. Kraft, J. Payandeh, G. Nakamura, J. T. Koerber, A. Rohou, *Science* **2020**, *367*, 1224.