

Angewandte Chemie www.angewandte.org

Post-Translational Modifications

How to cite: Angew. Chem. Int. Ed. 2022, 61, e202110223 doi.org/10.1002/anie.202110223 International Edition: German Edition: doi.org/10.1002/ange.202110223

Site-Selective Installation of N^{ε}-Modified Sidechains into Peptide and Protein Scaffolds via Visible-Light-Mediated Desulfurative C-C Bond **Formation**

Rhys C. Griffiths⁺, Frances R. Smith⁺, Jed E. Long, Daniel Scott, Huw E. L. Williams, Neil J. Oldham, Robert Layfield, and Nicholas J. Mitchell*

Abstract: Post-translational modifications (PTMs) enhance the repertoire of protein function and mediate or influence the activity of many cellular processes. The preparation of sitespecifically and homogeneously modified proteins, to apply as tools to understand the biological role of PTMs, is a challenging task. Herein, we describe a visible-light-mediated desulfurative $C(sp^3)-C(sp^3)$ bond forming reaction that enables the siteselective installation of N^{ε} -modified sidechains into peptides and proteins of interest. Rapid, operationally simple, and tolerant to ambient atmosphere, we demonstrate the installation of a range of lysine (Lys) PTMs into model peptide systems and showcase the potential of this technology by siteselectively installing an $N^{\varepsilon}Ac$ sidechain into recombinantly expressed ubiquitin (Ub).

 ${oldsymbol{P}}$ ost-translational modifications (PTMs) play a crucial role in protein function, controlling a myriad of biological pathways that influence normal cell biology and disease progression.^[1] The installation of PTMs, a process largely regulated by enzymes, occurs at many of the canonical amino acids.^[2] Such modifications significantly expand the diversity of chemical groups available across the proteome resulting in a considerable variety of protein isoforms that govern precise control over biological systems. The amino acid lysine (Lys) is known to display a particularly broad range of PTMs including methylation,^[3] acetylation,^[4] propionylation/butyrylation,^[5] succinvlation,^[6] formylation,^[7] benzoylation,^[8] β-

[*]	Dr. R. C. Griffiths, ^[+] Dr. F. R. Smith, ^[+] Prof. N. J. Oldham, Dr. N. I. Mitchell
	School of Chemistry, University of Nottingham
	University Park, Nottingham, NG7 2RD (UK)
	E-mail: nicholas.mitchell@nottingham.ac.uk
	Dr. J. E. Long, Dr. H. E. L. Williams
	Biodiscovery Institute, University of Nottingham
	University Park, Nottingham, NG7 2RD (UK)
	Dr. D. Scott, Prof. R. Layfield
	School of Life Sciences, Queen's Medical Centre
	University of Nottingham, Nottingham, NG7 2UH (UK)
[†]	These authors contributed equally to this work.
	Supporting information and the ORCID identification number(s) for
D	the author(s) of this article can be found under:
~	https://doi.org/10.1002/anie.202110223.
0	© 2021 The Authors. Angewandte Chemie International Edition
	published by Wiley-VCH GmbH. This is an open access article under
	the terms of the Creative Commons Attribution License, which

hydroxybutylation,^[9] and several others.^[10] These modifications mediate processes such as DNA packaging, gene expression, metabolism, and protein degradation.^[10] Despite significant research within this field, the function of the entire library of Lys PTMs remains largely unknown. It is therefore necessary to prepare natively modified proteins to better understand the molecular mechanisms in which they are involved. Thus, techniques that allow the generation of sitespecifically modified proteins are in high demand.

Non-canonical amino acids can be incorporated into proteins via the technique of amber codon suppression.^[11] While this strategy is undeniably powerful, it requires customized tRNA and extensive mutation of the relevant enzymatic machinery. Due to the complexities of this technique, synthetic approaches have also been sought. The total chemical synthesis of modified proteins via peptide ligation using native chemical ligation^[12,13] (NCL), expressed protein ligation (EPL),^[14,15] and various advances to these methods,^[16-20] has significantly contributed to the study of PTMs. However, a more accessible approach involves the preparation of a library of modified proteins from recombinantly expressed material via site-selective modification of specific residues. Such techniques must be chemoselective in the presence of the diverse chemical functionality displayed by the 20 canonical amino acids. This selectivity must also be effective under mild conditions i.e., in an aqueous environment at neutral pH, under ambient temperature and atmosphere. Owing to the superior nucleophilicity of the thiol group of cysteine (Cys), and its relatively low abundance across the eukaryotic proteome (ca. 2%), several popular techniques exploit this residue to selectively install PTMs and PTM mimics.

Utilizing appropriate reagents, mimics of several PTMs have been installed at Cys residues via nucleophilic substitution^[21] and thiol-ene^[22] chemistry (Figure 1 A), including Lys methylation and acetylation. Conversion of Cys residues to dehydroalanine (Dha) has also been explored to install Lys PTM mimics via Michael addition, with loss of native stereochemistry at the modified residue.^[23] While such mimics do enable biology mediated by PTMs to be investigated, methods that result in the formation of the native modification are preferred. Recent developments employing free-radical-addition to Dha residues enable $C(sp^3)-C(sp^3)$ bond forming chemistry to be achieved to install PTMs (Figure 1B).^[24-26] This technique is versatile and powerful; however, the reaction is not stereoselective, and the chemistry

Angew. Chem. Int. Ed. 2022, 61, e202110223 (1 of 6)

on under permits use, distribution and reproduction in any medium, provided the original work is properly cited.



Figure 1. Synthetic approaches to the site-selective installation of Lys PTM mimics (A) and PTMs (B); N^{e} Ac shown. Our approach (C): site-selective installation of N^{e} -modified sidechains via interception of visible-light-mediated desulfurization.

must be carried out under an inert atmosphere, necessitating use of a glove box and degassing of all solutions.

Previous work within our group describes a site-selective method for the modification of peptides and proteins via interception of free-radical-mediated dechalcogenation.[27] This reaction proceeds via trapping of an alanyl-radical intermediate produced during phosphine-mediated desulfurization/deselenization of Cys/selenocysteine (Sec) using a bespoke persistent radical trap with retention of stereochemistry at the modified residue. Further to this study, we postulated that by employing appropriately modified allylamine as the trap, this approach could be exploited to selectively install the N^{ε} -modified sidechain of Lys via C(sp³)-C(sp³) bond formation under ambient atmosphere without eliminating the stereochemistry at the modified residue (Figure 1 C). Addition of a primary alkyl radical to an unactivated alkene is a significant synthetic challenge. Freeradical-mediated C-C bond formation relies on an initial stabilized radical and/or addition onto an activated alkene.[28] Due to the necessary constraints imposed upon our approach to enable the installation of native chemistry, limiting H-atom abstraction by the intermediate alanyl radical was identified as our main obstacle. However, within the context of polypeptide bioconjugation at milligram/microgram scale, a large excess of the alkene can practically be employed to push the reaction towards completion in an attempt to outcompete the undesired pathway.

To explore this concept, peptide 1 (Ac-CWHISKEY-NH₂, a model carrying the majority of proteinogenic chemical functionality) was treated to desulfurization conditions^[29,30] in the presence of a large excess of N-allyl acetamide. At a concentration of 1 mM (1 dissolved into 20% dimethylsulf-(DMSO) in 6 M guanidinium hydrochloride oxide (Gdn·HCl), 0.1 M Na₂HPO₄, pH 7), the peptide was irradiated under blue light in the presence of an iridium(III) photocatalyst ((Ir[dF(CF₃)ppy]₂(dtbpy))PF₆, 1 mol%) to generate a thiyl radical. The water-soluble phosphine, TCEP (tris(2-carboxyethyl)-phosphine; 50 mM), was employed to facilitate desulfurization via a phosphoranyl radical, and Nallyl acetamide (2) (500 mM) was introduced to trap the intermediate alanyl radical. Irradiation was achieved using a simple apparatus consisting of inexpensive blue LEDs wrapped around a Pyrex dish on a stirrer plate (photochemistry set up 1, Supporting Information). Encouragingly, we observed full consumption of the starting peptide in under 2 hours to afford product 3a with the mass of the desired conjugate at 36% conversion (Table 1, entry 1). The alanine (Ala)-containing by-product, **3b** (Ac-AWHISKEY-NH₂), produced through trapping of the alanyl radical via H-atom abstraction, accounted for the majority of the remaining material (with a small percentage of a multiple addition product detected).

To optimize the reaction, the relevant variables were systematically evaluated (Table 1 and Supporting Information). The reaction proceeds only in the presence of the photocatalyst under irradiation of blue light (450-495 nm). Decreasing the peptide concentration to 0.5 mM improved the conversion (entry 2). Decreasing the TCEP concentration to 2.5 mM was not observed to hinder the reaction (entry 3). Employing an alternative photocatalyst (eosin Y) did not improve the conversion (Supporting Information). The optimal pH was found to be 8 (entry 4); increasing the iridium catalyst to 5 mol% decreased the time to completion to 60 mins with no change in yield (entry 5). Reducing the concentration of N-allyl acetamide to 100 mM did not hinder to the reaction (entry 6); below this threshold the reaction conversion decreased (Supporting Information). The % conversion to product was comparable when the reaction was run in buffer with no organic component (Supporting Information). The optimal conditions for this model appeared to be entry 6 (referred to hereafter as protocol A), affording 55% conversion by analytical HPLC. A preparative-scale reaction was run using this protocol and the progress monitored closely by HPLC. The starting peptide was consumed over 45 mins and the desired product 3a was isolated in 51% yield. Due to the use of an unactivated alkene, we were unable to fully out-complete H-atom abstraction. Under the optimized conditions, the Ala byproduct accounts for less than 40% of the remaining material, while multiple addition of N-allyl acetamide accounts for less than 10%. The reaction did not require degassing and could be run "on the bench" under ambient atmosphere. Perform-

Table 1: Optimization of visible-light-mediated installation of N^eAc sidechains; iridium(III) catalyst, (Ir[dF(CF₃)ppy]₂(dtbpy))PF₆, shown.



[a] 6 M Gdn·HCl, 0.1 M Na₂HPO₄, organic component DMSO, MeOH or MeCN. [b] % conversion calculated by analytical HPLC. [c] Isolated yield given in brackets. [d] All reactions conducted under blue LED lights at room temp. (set up 1), except for entry 7. [e] PhotoRedOx Box used for entry 7 (set up 2); 34 mWcm⁻² LED bulb, 450 nm.

ing the reaction using degassed solvents in a glove box did not improve conversion. Protocol A was then repeated in a temperature controlled PhotoRedOx Box (HepatoChem) using a 34 mW cm⁻² LED bulb at 450 nm (set up 2, Supporting Information). Remarkably, the reaction reached completion within just 2 mins using this apparatus, in comparable conversion and isolated yield (entry 7). Finally, increasing the mol% of the catalyst above 5 mol% did not improve the conversion for this model (entry 8).

To ensure the stereochemistry at the modified residue was not compromised during the reaction, model peptides Ac-CAY-NH₂ (**4a**) and Ac-D-CAY-NH₂ (**4b**) were synthesized and subjected to protocol A, yielding the desired epimers **5a** and **5b**. Comparison of the NMR spectra of these products with the NMR of **5a** synthesized on the solid phase confirmed that the integrity of the α -stereocentre of the target residue remains intact throughout the modification process (Figure 2). Furthermore, to explore the installation of multiple modifications, peptide **6a** (Ac-YECPLAHISCKY-NH₂) was prepared and subjected to protocol A using trap **2**. As expected, several products were identified as the various permutations of the N^eAc sidechain and Ala at the two Cys sites (Supporting Information) indicating that this method is only suitable for the installation of single-point modifications.

To further probe the effect of the local chemical environment on the reaction, a small selection of tetrapeptides was prepared and subjected to protocol A using trap 2 (Ac-XCAY-NH₂; X = Lys (7), Glu (glutamic acid; 8), Tyr (tyrosine; 9), Ile (isoleucine; 10)). These models carry an H-atom donor (phenol), charged groups at neutral pH (ammonium and carboxylate groups), and steric encumbrance directly adjacent to the target Cys. The extent of Ala by-product formation and multiple addition varied slightly across this series, however the generation of the desired products (11–14) remained broadly consistent with the previous model (see Table 2). Co-elution of product **11** with impurities impeded isolation of the desired product for this example.

To investigate the site-selective installation of a range of N^{ε} -modified sidechains, several modified allyl compounds



Figure 2. NMR spectra (CH α region) of N^eAc sidechain installation using model peptides **4a** and **4b** to afford epimers **5a** (B) and **5b** (A); NMR spectra of **5a** synthesised on the solid phase (C).

Table 2: Effect of H-atom donor, electrostatics, and steric encumbrance on the efficiency of the reaction.





were synthesized starting from allyl amine or allyl chloride to afford a small library of traps (2, 15–22) (Figure 3; see Supporting Information for experimental details). Using our method, these traps enable installation of a range of Lys PTMs into polypeptides including acetylation (2), methylation (15–17), formylation (18), propionylation (19), β hydroxybutylation (20), succinylation (21), and benzoylation



(22). A 10 mer sequence representing residues 32-41 of histone H2A (homo sapiens) (23) was synthesized as a biologically relevant model to explore the reaction scope using these traps. Protocol A conditions were applied using trap 2 to afford the desired product 24 in 50% conversion (by analytical HPLC) over 3 hours using set up 1. Slight alterations to this protocol (i.e., an increase in the excess of TCEP from 2.5 mM to 5 mM and the catalyst from 5 mol% to 10 mol%) afforded a higher conversion for this model over 1 hour (62%-referred to as protocol B). The protocol B conditions were then applied on a preparative scale to model peptide 23 with traps 2, 15-22 in our simple photochemistry set up 1 to afford a selection of peptides bearing an internally modified Lys residue (24-32) in moderate-good yield (47-62%), with full consumption of 23 in under 60 mins (Figure 4). These substrate-scope reactions were then repeated in the PhotoRedOx Box (set up 2) using protocol B, for comparison. As observed previously, the reactions reached completion within 2 mins to afford the desired products 24-32 in comparable isolated yield to those run using set up 1 (48-68%; Figure 4). Furthermore, to explore the installation of a Lys PTM into a cyclic peptide, model 33 (c[WHISCKEY]) was synthesized and subjected to protocol



Figure 4. Site-selective installation of a range of Lys PTMs via visible-light-mediated, desulfurative C(sp³)–C(sp³) bond formation. [a] Photochemical set-up 1 (blue LEDs). [b] Photochemical set-up 2 (PhotoRedOx Box); isolated yields shown.

Angew. Chem. Int. Ed. 2022, 61, e202110223 (4 of 6)

© 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

B (set up 2) using probe **16** to afford peptide **34** in 46% isolated yield (see Supporting Information).

Satisfied that this method enabled effective installation of a range of PTMs in model peptide systems, we applied our reaction to investigate the incorporation of a biologically relevant PTM using recombinantly expressed protein material (Figure 5). Acetylation of ubiquitin (Ub) at Lys48 (K48), regulates polyubiquitination of protein substrates and thus controls proteasome-mediated protein degradation.^[31] Employing a K48C mutant of Ub (**35**) (the native Ub



Figure 5. Site-selective installation of an N^{ϵ} Ac sidechain into recombinantly expressed Ub **35** (K48C mutant).

sequence contains no Cys residues, see Supporting Information), produced via recombinant expression in *E. coli*,^[32] we subjected the material to protocol B in the presence of *N*-allyl acetamide (**2**) in set up 1 to demonstrate the accessibility of our technique using inexpensive apparatus. Sub-optimal conversion to the product prompted us to increase the concentration of TCEP described in protocol B to 50 mM (referred to as protocol C).

Gratifyingly, the reaction proceeded to the expected conversion over 60 mins using these conditions. The desired product 36 was isolated by preparative HPLC in 44% yield (Figure 5). Considering the size and chemical complexity of Ub, the operational simplicity of the protocol, and the application of this method to the installation of native modifications via C(sp3)-C(sp3) bond formation, we consider the yield of 36 to be acceptable. To ensure the modified Ub could be re-folded from the denaturing conditions of the bioconjugation, product 36 was re-dissolved in 6 M Gdn·HCl, 0.1 M Na₂HPO₄, pH 7, folded via dialysis into 25 mM Na₂HPO₄, 100 mM NaCl, pH 7.0, and analyzed via ¹H NMR. The extended NH region of the spectra (6.5-9.5 ppm), and upshifted methyl signals indicating hydrophobic packing, were consistent with folded WT Ub.^[33] The modified N^eAc-Ub 36 protein retained integrity, based on migration on an SDS-PAGE gel, and comparable recognition to K48C-Ub was observed via western blot using an antiubiquitin antibody (Supporting Information).

In conclusion, we have developed a rapid and operationally simple reaction that enables the site-selective installation of Lys PTMs into complex polypeptides via $C(sp^3)-C(sp^3)$ bond formation. The reaction proceeds via visible-lightmediated desulfurization under ambient conditions in the presence of an unactivated alkene and is tolerant to a range of modifications. The optimized protocol translates well to a small protein system, enabling isolation of modified material in acceptable yield. To our knowledge, this report represents the first example of direct, late-stage installation of Lys PTMs via chemical conversion of a canonical residue without disrupting the stereochemistry of the target amino acid. Due to the operational simplicity of the reaction, and inexpensive experimental set-up, the described protocols are accessible to researchers without extensive training in synthetic chemistry to enable the production of biochemical tools to further expand our knowledge of the chemically modified proteome.

Acknowledgements

This work was supported by funding from the Engineering and Physical Sciences Research Council (EP/S028323/1, EP/ S017739/1) and from the University of Nottingham (PhD studentship for R.C.G.). We gratefully acknowledge Ben Pointer-Gleadhill for assistance with MS analysis.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bioconjugation · cysteine · peptides · photochemistry · site selectivity

- R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, *Nat. Chem. Biol.* **2018**, *14*, 206–214.
- [2] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto Jr, Angew. Chem. Int. Ed. 2005, 44, 7342–7372; Angew. Chem. 2005, 117, 7508–7539.
- [3] M. Luo, Chem. Rev. 2018, 118, 6656-6705.
- [4] I. Ali, R. J. Conrad, E. Verdin, M. Ott, Chem. Rev. 2018, 118, 1216–1252.
- [5] Y. Chen, R. Sprung, Y. Tang, H. Ball, B. Sangras, S. C. Kim, J. R. Falck, J. Peng, W. Gu, Y. Zhao, *Mol. Cell. Proteomics* 2007, 6, 812–819.
- [6] Z. Zhang, M. Tan, Z. Xie, L. Dai, Y. Chen, Y. Zhao, Nat. Chem. Biol. 2011, 7, 58–63.
- [7] T. Jiang, X. Zhou, K. Taghizadeh, M. Dong, P. C. Dedon, Proc. Natl. Acad. Sci. USA 2007, 104, 60–65.
- [8] H. Huang, D. Zhang, Y. Wang, M. Perez-Neut, Z. Han, Y. G. Zheng, Q. Hao, Y. Zhao, *Nat. Commun.* 2018, 9, 3374.
- [9] Z. Xie, D. Zhang, D. Chung, Z. Tang, H. Huang, L. Dai, S. Qi, J. Li, G. Colak, Y. Chen, *Mol. Cell* **2016**, *62*, 194–206.
- [10] Z. A. Wang, P. A. Cole, Cell Chem. Biol. 2020, 27, 953-969.
- [11] H. Neumann, S. Y. Peak-Chew, J. W. Chin, Nat. Chem. Biol. 2008, 4, 232–234.
- [12] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. Kent, Science 1994, 266, 776–779.
- [13] S. B. Kent, Chem. Soc. Rev. 2009, 38, 338-351.
- [14] M. M. Müller, T. W. Muir, Chem. Rev. 2015, 115, 2296-2349.
- [15] T. W. Muir, D. Sondhi, P. A. Cole, Proc. Natl. Acad. Sci. USA 1998, 95, 6705–6710.
- [16] N. J. Mitchell, L. R. Malins, X. Liu, R. E. Thompson, B. Chan, L. Radom, R. J. Payne, J. Am. Chem. Soc. 2015, 137, 14011–14014.



- [17] N. J. Mitchell, J. Sayers, S. S. Kulkarni, D. Clayton, A. M. Goldys, J. Ripoll-Rozada, P. J. B. Pereira, B. Chan, L. Radom, R. J. Payne, *Chem* **2017**, *2*, 703–715.
- [18] J. Sayers, P. M. Karpati, N. J. Mitchell, A. M. Goldys, S. M. Kwong, N. Firth, B. Chan, R. J. Payne, *J. Am. Chem. Soc.* 2018, 140, 13327-13334.
- [19] A. C. Conibear, E. E. Watson, R. J. Payne, C. F. Becker, *Chem. Soc. Rev.* 2018, 47, 9046–9068.
- [20] S. S. Kulkarni, J. Sayers, B. Premdjee, R. J. Payne, Nat. Rev. Chem. 2018, 2, 0122.
- [21] M. D. Simon, F. Chu, L. R. Racki, C. Cecile, A. L. Burlingame, B. Panning, G. J. Narlikar, K. M. Shokat, *Cell* **2007**, *128*, 1003 – 1012.
- [22] F. Li, A. Allahverdi, R. Yang, G. B. J. Lua, X. Zhang, Y. Cao, N. Korolev, L. Nordenskiöld, C. F. Liu, *Angew. Chem. Int. Ed.* **2011**, 50, 9611–9614; *Angew. Chem.* **2011**, *123*, 9785–9788.
- [23] J. Guo, J. Wang, J. S. Lee, P. G. Schultz, Angew. Chem. Int. Ed. 2008, 47, 6399-6401; Angew. Chem. 2008, 120, 6499-6501.
- [24] T. H. Wright, B. J. Bower, J. M. Chalker, G. J. Bernardes, R. Wiewiora, W.-L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée, A. Phanumartwiwath, *Science* 2016, 354, aag1465.
- [25] B. Josephson, C. Fehl, P. G. Isenegger, S. Nadal, T. H. Wright, A. W. Poh, B. J. Bower, A. M. Giltrap, L. Chen, C. Batchelor-McAuley, *Nature* 2020, 585, 530-537.

- [26] A. Yang, S. Ha, J. Ahn, R. Kim, S. Kim, Y. Lee, J. Kim, D. Söll, H.-Y. Lee, H.-S. Park, *Science* 2016, 354, 623–626.
- [27] R. C. Griffiths, F. R. Smith, J. E. Long, H. E. Williams, R. Layfield, N. J. Mitchell, *Angew. Chem. Int. Ed.* **2020**, *59*, 23659–23667; *Angew. Chem.* **2020**, *132*, 23867–23875.
- [28] B. Giese, Angew. Chem. Int. Ed. Engl. 1983, 22, 753-764; Angew. Chem. 1983, 95, 771-782.
- [29] Q. Wan, S. J. Danishefsky, Angew. Chem. Int. Ed. 2007, 46, 9248– 9252; Angew. Chem. 2007, 119, 9408–9412.
- [30] X.-F. Gao, J.-J. Du, Z. Liu, J. Guo, Org. Lett. 2016, 18, 1166– 1169.
- [31] F. Ohtake, Y. Saeki, K. Sakamoto, K. Ohtake, H. Nishikawa, H. Tsuchiya, T. Ohta, K. Tanaka, J. Kanno, *EMBO Rep.* 2015, 16, 192–201.
- [32] T. P. Garner, J. Strachan, E. C. Shedden, J. E. Long, J. R. Cavey, B. Shaw, R. Layfield, M. S. Searle, *Biochemistry* 2011, 50, 9076– 9087.
- [33] Biological Magnetic Resonance Data Bank (BMRDB); entries 68 and 4769; http://www.bmrb.wisc.edu/.

Manuscript received: July 30, 2021

- Revised manuscript received: October 13, 2021
- Accepted manuscript online: October 29, 2021
- Version of record online: December 3, 2021