ChemBioChem

Research Article
doi.org/10.1002/cbic.202200258

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

Special Collection

Context-Dependence of the Reactivity of Cysteine and Lysine Residues

Linus B. Boll[a] and Ronald T. Raines*[a]

Dedicated to the memory of Ulf Diederichsen

The *S*-alkylation of Cys residues with a maleimide and the N^{ϵ}-acylation of Lys residues with an *N*-hydroxysuccinimide (NHS) ester are common methods for bioconjugation. Using Cys and Lys derivatives as proxies, we assessed differences in reactivity depending on the position of Cys or Lys in a protein sequence. We find that Cys position is exploitable to improve site-selectivity in maleimide-based modifications. Reactivity decreases substantially in the order N-terminal > in-chain > C-terminal Cys due to modulation of sulfhydryl pK_a by the α-ammonium and carboxylate groups at the termini. A lower pK_a value yields a larger fraction thiolate, which promotes selectivity while somewhat decreasing thiolate nucleophilicity in accord with $\beta_{nuc} = 0.41$. Lowering pH and salt concentration enhances selectivity still further. In contrast, differences in the reactivity of Lys towards an NHS ester were modest due to an appreciable decrease in amino group nucleophilicity with a lower pK_a of its conjugate acid. Hence, site-selective Lys modification protocols will require electrophiles other than NHS esters.

The targeting of Cys side chains with maleimides and Lys side chains with *N*-hydroxysuccinimide (NHS) esters dominates the landscape of residue-specific modification of proteins.[1–4] Proteins modified thusly are used as antibody-drug conjugates[5,6] and in numerous other applications.[1–4]

The low abundance of Cys residues (2.4% in human proteins[7]) and selectivity of maleimides can allow for single residues to be modified, whereas the higher abundance of Lys residues (5.0% in human proteins[7]) and possible modification of the N-terminus or other side chains by NHS esters pose a challenge to site-selectivity.[4–6,8] Three attributes of the protein microenvironment also contribute to selectivity: (1) solvent accessibility, (2) intrinsic affinity for the electrophile, and (3) perturbation of the pK_a value of the nucleophile.[9–14] Whereas the first two attributes are idiosyncratic to a particular protein, the last enables general conclusions and is the focus of our work.

Covalent modification of a Cys or Lys residue requires an anionic thiolate[15] or neutral amine, respectively. Accordingly, we reasoned that N-terminal, in-chain or C-terminal Cys and Lys residues should react at different rates based on proximity to the cationic α-ammonium group at the N-terminus or anionic carboxylate group at the C-terminus. To our knowledge, these reactivity differences, which could be leveraged to enhance site-selectivity, have not yet been evaluated systematically. Here, we use amino acid derivatives as proxies to discern how the reactivity of a Cys side chain towards a maleimide and a Lys side chain towards an NHS ester depend on their position in the sequence of a protein or peptide. Our findings highlight the context-dependence of Cys and Lys reactivity and provide guidance on the selectivity achievable from pK_a alone.

As models for an N-terminal, in-chain, or C-terminal Cys residue, we chose H–Cys–NH_2 (**1**), Ac–Cys–NH_2 (**2**), and Ac–Cys–OH (**3**), respectively. Thiol **1** was prepared by deprotection of H–Cys(Trt)–NH_2. Thiol **2** was synthesized by acetylation on Rink amide resin. *N*-Ethylmaleimide (NEM) was chosen as a representative maleimide.

Relative reactivities of thiols **1**–**3** towards NEM in PBS (pH 7.4) were assessed in competition experiment **I**, which featured equimolar amounts of each thiol and NEM (Figure 1a). Incubation for < 5 min sufficed to consume the NEM. This short time-course averted complexity from the comparatively slow degradation of the thiosuccinimide products by hydrolysis or intramolecular transcyclization.[16,17] RP-HPLC traces (Figures S5–S7) revealed two peaks for each product due to the formation of diastereomers.[17] Linear response curves for each Michael addition product in the range of 0.05–1.0 mM were determined for calibration (Figures S1–S4).

Thiols **1**–**3** differed significantly in their reactivity (Figure 1b). Thiol **1** was the most reactive towards NEM, and thiol **3** was the least. We reasoned that these differences could arise from different pK_a values, as the sulfhydryl group of thiol **1** is proximal to a cationic ammonium group whereas that of thiol **3** is proximal to an anionic carboxylate group.

To test this hypothesis, we measured the sulfhydryl pK_a values of thiols **1**–**3** by UV spectroscopy, exploiting the increase of A_{238} upon deprotonation.[18–20] We found that thiol **2** has pK_a = 8.88(1) (Figure 2). Thiol **3** is less acidic with pK_a = 9.45(3), which is indistinguishable from the value obtained with potentiometric titrations.[21,22] The sulfhydryl group of thiol **1** has microscopic pK_a values that depend on the protonation state of

[a] *L. B. Boll, Prof. Dr. R. T. Raines*
Department of Chemistry, Massachusetts Institute of Technology
Cambridge, Massachusetts 02139 (USA)
E-mail: rtraines@mit.edu

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202200258

This article belongs to a Joint Special Collection dedicated to Ulf Diederichsen.

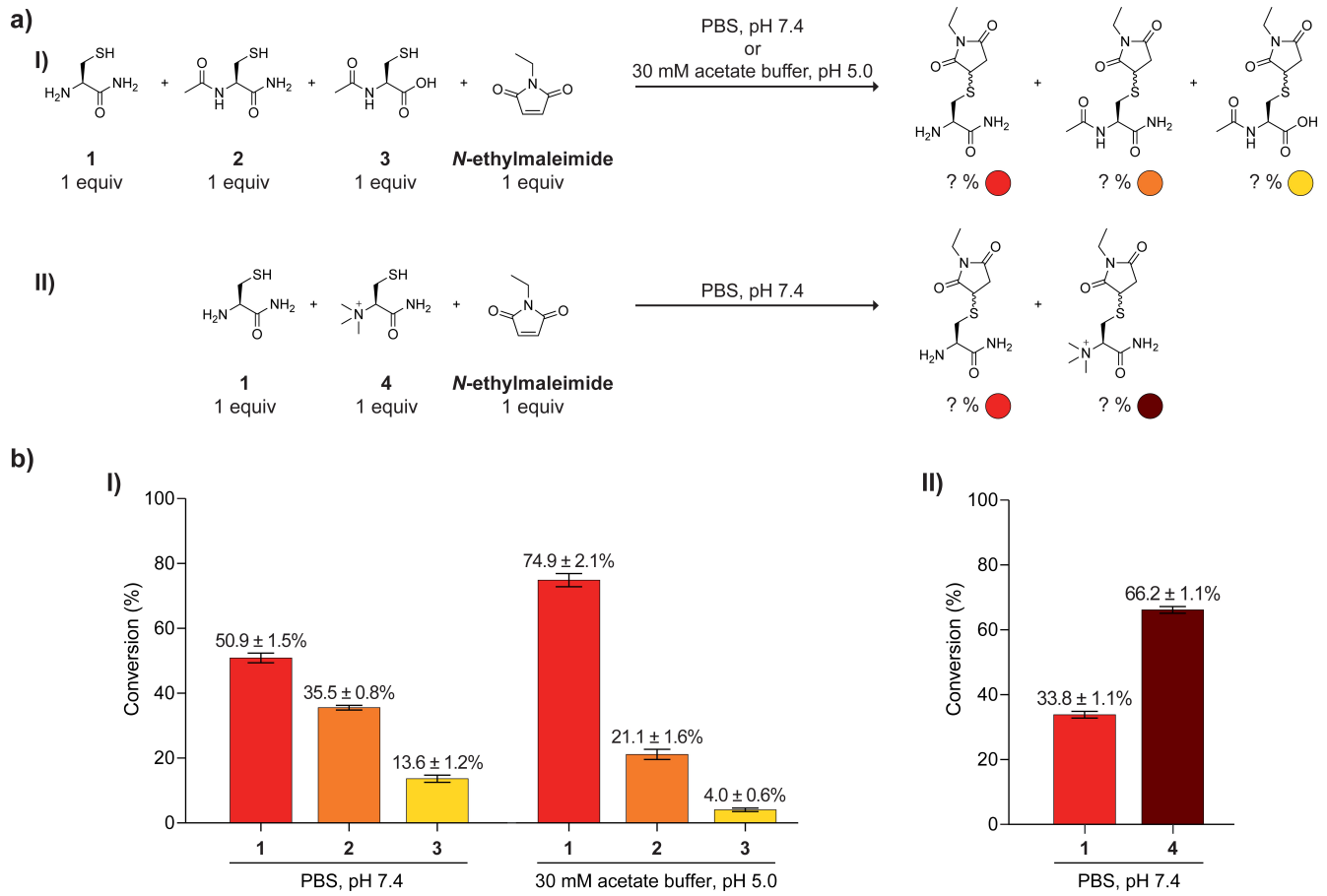
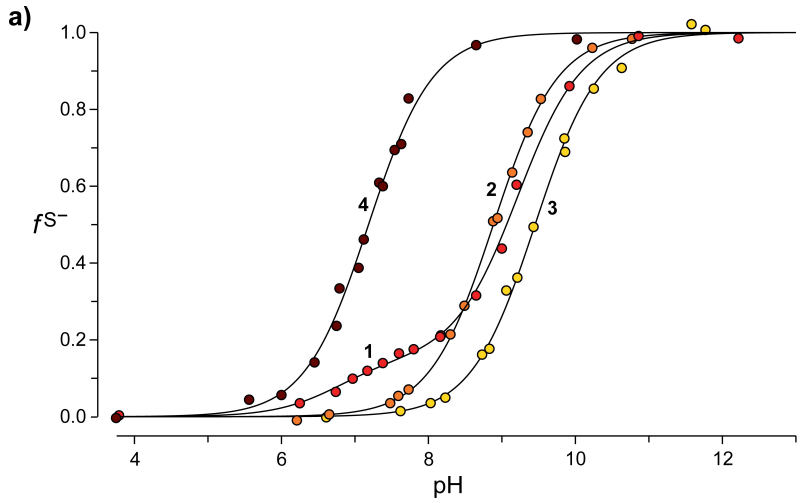


Figure 1. a) Competition experiments featuring thiols **1**–**4** set up to determine relative reactivities. b) Product distributions determined for each assay. Percentages reflect shares of products with respect to the total amount of Michael addition products.

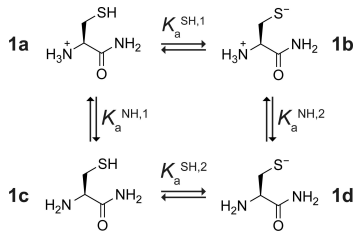


b)

Thiol	Sulfhydryl pK_a	f^{S-} (pH 7.4)
H-Cys-NH_2 (**1**,●)	7.57(20), 9.15(3)	0.135
Ac-Cys-NH_2 (**2**,●)	8.88(1)	0.032
Ac-Cys-OH (**3**,●)	9.45(3)	0.009
$(Me)_3$-Cys-NH_2 (**4**,●)	7.19(4)	0.621

Figure 2. Determination of sulfhydryl pK_a values by UV titration monitoring A_{238} relative to A_{238} of the buffer containing no thiol. a) Fraction thiolate (f^{S-}), calculated as a function of pH. Data were fitted to a single titration model (thiols **2**–**4**) or to a titration model accounting for microscopic pK_a values (thiol **1**) (Equation S1). b) Table of measured sulfhydryl pK_a and calculated f^{S-} (pH 7.4) values.

its amino group,[23] giving four relevant species, **1a**–**1d** (Scheme 1). Applying a model that accounts for the interdependence of the sulfhydryl and ammonium pK_a values (Equation S1), we found that $pK_a^{SH,1} = 7.57(20)$ and $pK_a^{SH,2} = 9.15(3)$ (Figure 2). Thus, the sulfhydryl pK_a value of thiol **1** is highly dependent on the protonation state of its amino group. The



Scheme 1. Species of thiol **1** relevant to reactivity towards NEM at pH = 7.4 and the dissociation constants relating these species. $pK_a^{SH,1} = 7.57(20)$, $pK_a^{SH,2} = 9.15(3)$, $pK_a^{NH,1} = 7.31(24)$, and $pK_a^{NH,2} = 8.53(9)$.

low sulfhydryl pK_a value of species **1a** leads to a larger fraction thiolate, f^{S^-}, for thiol **1** than for thiol **2** (Figure 2), which likely manifests in higher reactivity towards NEM.

To solidify our interpretation of the reactivity of thiol **1**, we made the charge on its ammonium group permanent by synthesizing $(Me)_3$–Cys–NH_2 (**4**). Thiol **4** was obtained by treating H–Cys(Trt)–NH_2 with iodomethane and deprotecting the sulfhydryl group. In competition experiment **II**, we found that thiol **4** was more reactive than thiol **1** (Figure 1b). The higher reactivity is consistent with the large f^{S^-} of thiol **4** that results from its single low sulfhydryl pK_a value of 7.19(4) (Figure 2).

Only thiolate species are relevant for reactivity with NEM.[15] Accordingly, the differences in f^{S^-} values should lead to larger differences in reactivity between thiols **1**–**4** than those observed experimentally. The effect of f^{S^-} on reactivity is, however, moderated by the nucleophilicity of a thiolate species, which increases with the pK_a value of its sulfhydryl group.[24] This moderation was quantified by using a Brønsted-type linear free-energy relationship (Equation S5). The Brønsted coefficient, β_{nuc}, is indicative of the charge that develops in the transition state.[25] Bednar reported $\beta_{nuc} = 0.43$ for the addition of several alkyl thiolates to NEM.[15] Keillor and coworkers found $\beta_{nuc} = 0.40$ for the addition of alkyl thiolates to *N*-methylmaleimide.[26] Our data in PBS gives $\beta_{nuc} = 0.41(7)$ (Figure 3), which is consistent with limited electron transfer from thiolate to maleimide in the transition state. This value agrees with a thiolate being a "soft" nucleophile.[27]

The relative reactivity of thiol **1** is enhanced upon lowering of the pH of the medium (Figure 1b). In 30 mM sodium acetate buffer, pH 5.0, thiol **1** exhibits almost 4-fold higher reactivity than in-chain thiol **2**. The explanation likely arises from a shift in the equilibrium between species **1a** and **1c** towards **1a**, which is more reactive.

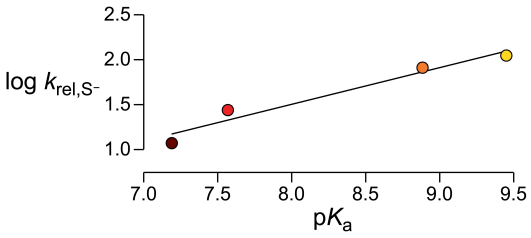


Figure 3. Brønsted plot of $\log(k_{rel,S^-})$ against pK_a, yielding $\beta_{nuc} = 0.41(7)$. For thiol **1**, $pK_a^{SH,1}$ was used for plotting. Thiolate concentration-independent relative rate constants k_{rel,S^-} were determined using Equation S4.

Salt concentration can affect pK_a values.[28–30] To assess the effect of salt concentration on the reactivity of Cys residues towards NEM, we repeated competition experiment **I** in buffers containing different levels of salt. In sodium acetate buffer, pH 5.0, the relative reactivity of thiols **1** and **3** was 21, 19, and 10, in 5, 30, and 300 mM buffer, respectively (Figure S8). A similar trend was observed in PBS, pH 7.4. These effects are subtle but indicate that higher selectivity for an N-terminal Cys residue is attainable at a lower salt concentration.

We probed the context-dependence of the reactivity of Lys residues in a similar manner. To begin, we installed two methyl groups on either the ε- or the α-amino group to isolate reactivity with minimal perturbation. We chose H–Lys$(Me)_2$–NH_2 (**5**) and $(Me)_2$–Lys–NH_2 (**6**) as proxies for an N-terminal Lys residue, Ac–Lys–NH_2 (**7**) for an in-chain position, and Ac–Lys–OH (**8**) for a C-terminal residue. We also used H–Ala–NH_2 (**9**) to examine the reactivity of an unperturbed N-terminal amino group. Amines **5**–**7** were synthesized by reductive methylation or acetylation on Rink amide resin. *N*-(Benzoyloxy)succinimide was chosen as a representative NHS ester.

Competition experiments **III**–**V** employed equimolar amounts of selected amines **5**–**9** and the NHS ester (Figure 4a) and were performed in PBS (pH 7.4). In aqueous buffer, the aminolysis of the NHS ester is competing with its hydrolysis.[31] Hence, additional peaks attributed to benzoic acid and *O*-acylated succinyl hydroxamic acid were observed (Figure S17) in the RP-HPLC chromatograms (Figures S14–S16). Linear response curves for each amidation product in the range of 0.05–1.0 mM were determined for calibration (Figures S9–S13).

Competition experiment **III** probed the relative reactivity of the amino groups in N-terminal Lys **5** and **6** and in-chain Lys **7**. The product distribution is indicative of similar reactivity for all three amino groups with a small preference towards acylation of the N-terminal Lys (Figure 4b). This enhanced reactivity is consistent with perturbations in pK_a values owing to a proximal ε- or α-ammonium group. The N-terminal α-amino group of **5** exhibits slightly higher reactivity than does the ε-amino group of **6**. This preference is expected given the lower average pK_a of an ammonium group at the N-terminus in comparison to one in a Lys side chain.[32,33] Nonetheless, the observed differences in reactivity are small, which likely results from an appreciable decrease in the intrinsic reactivity towards NHS esters of amino groups with lower ammonium pK_a values.[34,35] Applying a Brønsted-type linear free-energy relationship, Cline and Hanna found $\beta_{nuc} = 0.73$ for the reactivity of primary amino groups towards *N*-(4-nitrobenzoyloxy)succinimide in 1,4-dioxane and $\beta_{nuc} = 1.0$ for reactivity towards *N*-(4-methoxybenzoyloxy)succinimide in aqueous solutions.[34,35] The magnitude of β_{nuc} highlights that reactivity advantages owing to a larger fraction of neutral amino group are nearly negated, consistent with the breakdown of the tetrahedral intermediate being rate-determining for the aminolysis of NHS esters.[34,35]

In competition experiment **IV**, amine **5** was replaced with H-Ala-NH_2 (**9**), representing an unperturbed N-terminal amino group. Although the pK_a of the ammonium group of amine **9** is expected

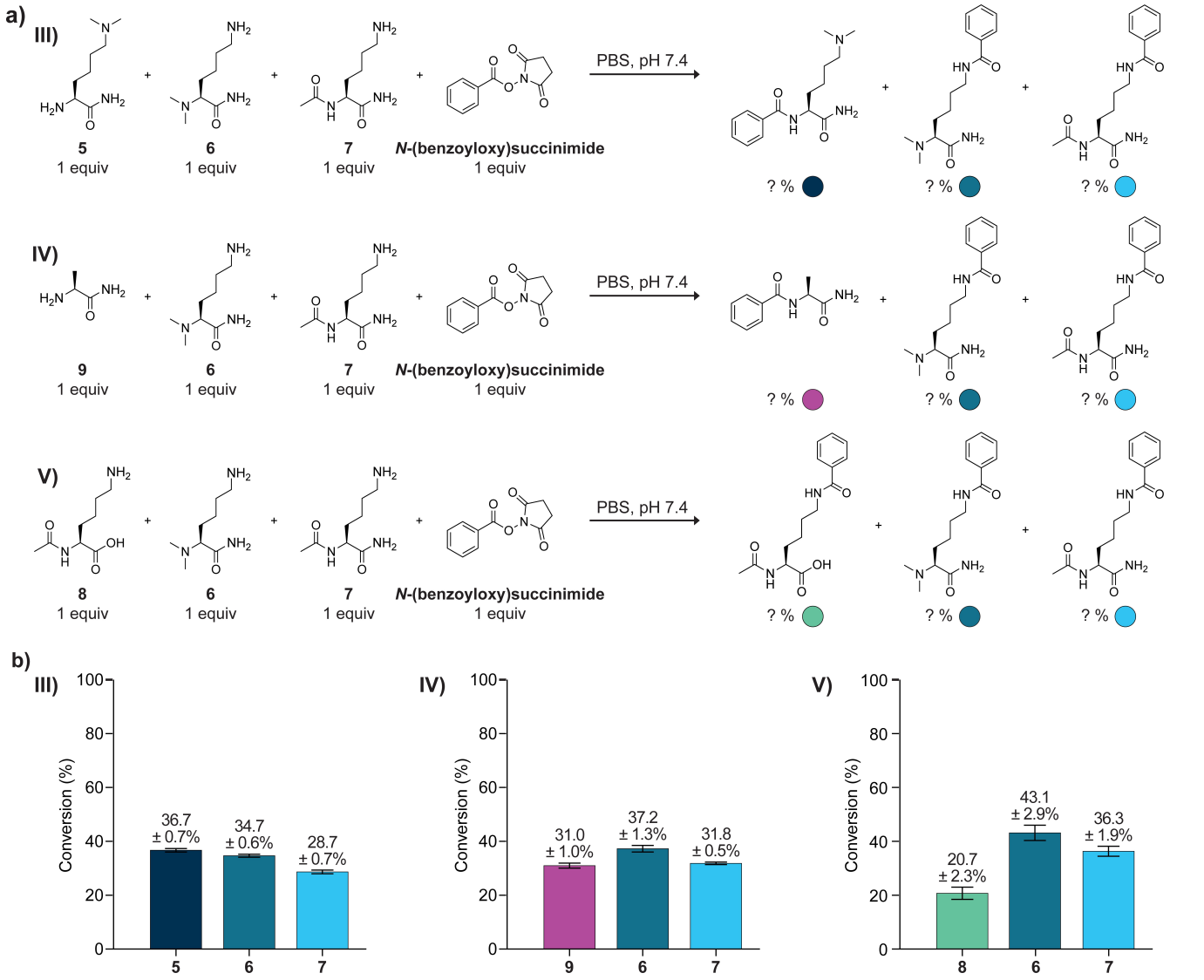


Figure 4. a) Competition experiments **III**–**V** involving amines **5**–**9** set up to determine relative reactivities. b) Product distributions determined for each assay. Percentages reflect shares of products with respect to the total amount of amidation products.

to be lower than that of amine **7**, their reactivity towards *N*-(benzoyloxy)succinimide is comparable (Figure 4b). We hypothesize the cause to be the branched carbon chain adjacent to the amino group of amine **9**, which introduces steric effects in the tetrahedral intermediate for *N*-acylation. Similarly, α-methylbenzylamine was found to have low reactivity towards *N*-(4-nitrobenzoyloxy)succinimide in comparison to unhindered amines with similar ammonium group pK_a values.[34] The sensitivity of aminolysis to steric effects could also explain why amine **5** is only slightly more reactive than amine **6**.

In competition experiment **V**, Ac–Lys–OH (**8**) was introduced as a proxy for a C-terminal Lys residue. Amine **8** has about half of the reactivity towards *N*-(benzoyloxy)succinimide in comparison to N-terminal amine **6** or in-chain amine **7** (Figure 4b). The negative charge on the carboxylate likely raises the pK_a value of the ammonium group of amine **8**, decreasing the concentration of neutral amine.

The relative reactivities observed herein suggest that only modest differences in reactivity of Lys residues or N-termini towards NHS esters should be expected based solely on their position in the sequence of a protein or peptide. Hence, it is not surprising that highly site-selective Lys modification by NHS esters remains challenging.[4] Particular Lys residues in a protein can exhibit heightened reactivity due to other factors, including elevated solvent accessibility, recognition of the acyl group attached to NHS, and large pK_a perturbations mediated by neighboring residues.[13,14,36] Attempts have been made to exploit such residues for labelling with NHS esters. Solvent accessibility studies enabled Adamo and coworkers to tune reaction conditions to modify Lys residues at three or six designated positions in CRM_{197}.[37] Weil and coworkers reported the selective modification of RNase A and lysozyme C at Lys1, but their analysis did not consider acylation of the more reactive N-terminal α-amino group.[38,39] As pointed out by Baker and coworkers, the carefully

tuned conditions and incomplete conversions required for homogenous modification will likely limit applications to proteins containing few Lys residues.[4] Thus, whereas the site-selective labelling of Lys residues featuring a unique protein microenvironment with NHS esters could be achievable,[4] current methods lack generality. Given the small differences in reactivity observed herein, a simple and broadly applicable yet site-selective method for amino group modification by an NHS ester is likely to be elusive.

In summary, we have used amino acid derivatives as proxies to investigate the context-dependence of Cys reactivity towards a maleimide and Lys reactivity towards an NHS ester. Competition assays revealed that reactivity decreases in the order N-terminal > in-chain > C-terminal Cys due to perturbation of sulfhydryl pK_a values by a proximal N-terminal α-ammonium group or a C-terminal carboxylate group. Selectivity for N-terminal modification can be enhanced by lowering the buffer pH and the salt concentration. The proximity of a Cys residue to cationic residues, the N-terminus of an α-helix, or the π* orbital of a main-chain carbonyl group can likewise decrease sulfhydryl pK_a values and increase reactivity.[40–42] Because of these attributes and the low abundance of Cys residues in natural proteins,[7] meaningful selectivity for its *S*-alkylation with a maleimide is achievable. In contrast, reactivity differences of N-terminal α-amino groups and Lys ε-amino groups towards an NHS ester were less pronounced. This dichotomy arises because the intrinsic reactivity of an amino group towards NHS esters decreases appreciably with lower ammonium pK_a of its conjugate acid. Hence, NHS ester-based protocols are disadvantageous if product homogeneity is important. These findings provide guidance for the planning and evaluation of bioconjugation experiments.

Acknowledgements

L.B.B. was supported by an Excellence Scholarship (ETH Zürich Foundation). This work was supported by Grant R01 GM044783 (NIH).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: acidity · amino acids · bioconjugation · maleimide · NHS ester

[1] G. T. Hermanson, *Bioconjugate Techniques*, 3rd ed., Academic Press, Waltham, **2013**.
[2] J. M. J. M. Ravasco, H. Faustino, A. Trindade, P. M. P. Gois, *Chem. Eur. J.* **2019**, *25*, 43–59.
[3] E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira, G. J. L. Bernardes, *Nat. Chem. Rev.* **2019**, *3*, 147–171.
[4] M. Haque, N. Forte, J. R. Baker, *Chem. Commun.* **2021**, *57*, 10689–10702.
[5] P. Akkapeddi, S.-A. Azizi, A. M. Freedy, P. M. S. D. Cal, P. M. P. Gois, G. J. L. Bernardes, *Chem. Sci.* **2016**, *7*, 2954–2963.
[6] A. Beck, L. Goetsch, C. Dumontet, N. Corvaïa, *Nat. Rev. Drug Discovery* **2017**, *16*, 315–337.
[7] N. Echols, P. Harrison, S. Balasubramanian, N. M. Luscombe, P. Bertone, Z. Zhang, M. Gerstein, *Nucleic Acids Res.* **2002**, *30*, 2515–2523.
[8] J. M. Chalker, G. J. L. Bernardes, Y. A. Lin, B. G. Davis, *Chem. Asian J.* **2009**, *4*, 630–640.
[9] M. P. Lutolf, N. Tirelli, S. Cerritelli, L. Cavalli, J. A. Hubbell, *Bioconjugate Chem.* **2001**, *12*, 1051–1056.
[10] E. S. Underbakke, Y. Zhu, L. L. Kiessling, *Angew. Chem. Int. Ed.* **2008**, *47*, 9677–9680; *Angew. Chem.* **2008**, *120*, 9823–9826.
[11] M.-C. Ho, J.-F. Ménétret, H. Tsuruta, K. N. Allen, *Nature* **2009**, *459*, 393–397.
[12] E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. D. Dillon, D. A. Bachovchin, K. Mowen, D. Baker, B. F. Cravatt, *Nature* **2010**, *468*, 790–795.
[13] S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia, B. F. Cravatt, *Nat. Chem.* **2017**, *9*, 1181–1190.
[14] M. E. Abbasov, M. E. Kavanagh, T.-A. Ichu, M. R. Lazear, Y. Tao, V. M. Crowley, C. W. am Ende, S. M. Hacker, J. Ho, M. M. Dix, R. Suciu, M. M. Hayward, L. L. Kiessling, B. F. Cravatt, *Nat. Chem.* **2021**, *13*, 1081–1092.
[15] R. A. Bednar, *Biochemistry* **1990**, *29*, 3684–3690.
[16] R. P. Lyon, J. R. Setter, T. D. Bovee, S. O. Doronina, J. H. Hunter, M. E. Anderson, C. L. Balasubramanian, S. M. Duniho, C. I. Leiske, F. Li, P. D. Senter, *Nat. Biotechnol.* **2014**, *32*, 1059–1062.
[17] M. Lahnsteiner, A. Kastner, J. Mayr, A. Roller, B. K. Keppler, C. R. Kowol, *Chem. Eur. J.* **2020**, *26*, 15867–15870.
[18] R. E. Benesch, R. Benesch, *J. Am. Chem. Soc.* **1955**, *77*, 5877–5881.
[19] K. J. Woycechowsky, K. D. Wittrup, R. T. Raines, *Chem. Biol.* **1999**, *6*, 871–879.
[20] J. C. Lukesh III, M. J. Palte, R. T. Raines, *J. Am. Chem. Soc.* **2012**, *134*, 4057–4059.
[21] B. Shen, C. Bazin, A. M. English, *J. Pept. Sci.* **2013**, *19*, 263–267.
[22] A. E. Fazary, N. S. Awwad, H. A. Ibrahium, A. A. Shati, M. Y. Alfaifi, Y.-H. Ju, *ACS Omega* **2020**, *5*, 19598–19605.
[23] O. R. Maguire, J. Zhu, W. D. G. Brittain, A. S. Hudson, S. L. Cobb, A. C. O'Donoghue, *Chem. Commun.* **2020**, *56*, 6114–6117.
[24] P. Nagy, *Antioxid. Redox Signaling* **2013**, *18*, 1623–1641.
[25] A. Williams, *Free Energy Relationships in Organic and Bio-organic Chemistry*, Royal Society of Chemistry, Cambridge, **2003**.
[26] M. A. R. Raycroft, K. É. Racine, C. N. Rowley, J. W. Keillor, *J. Org. Chem.* **2018**, *83*, 11674–11685.
[27] R. G. Pearson, *Chemical Hardness: Applications from Molecules to Solids*, Wiley-VCH, Weinheim, **1997**.
[28] S. R. Milner, *Phil. Mag.* **1912**, *23*, 551–578.
[29] S. R. Milner, *Phil. Mag.* **1913**, *25*, 742–751.
[30] P. Debye, E. Hückel, *Phys. Z.* **1923**, *24*, 185–206.
[31] O. Klykov, M. G. Weller, *Anal. Methods* **2015**, *7*, 6443–6448.
[32] G. R. Grimsley, J. M. Scholtz, C. N. Pace, *Protein Sci.* **2009**, *18*, 247–251.
[33] S. Pahari, L. Sun, E. Alexov, *Database* **2019**, *2019*, baz024.
[34] G. W. Cline, S. B. Hanna, *J. Am. Chem. Soc.* **1987**, *109*, 3087–3091.
[35] G. W. Cline, S. B. Hanna, *J. Org. Chem.* **1988**, *53*, 3583–3586.
[36] C. C. Ward, J. I. Kleinman, D. K. Nomura, *ACS Chem. Biol.* **2017**, *12*, 1478–1483.
[37] S. Crotti, H. Zhai, J. Zhou, M. Allan, D. Proietti, W. Pansegrau, Q.-Y. Hu, F. Berti, R. Adamo, *ChemBioChem* **2014**, *15*, 836–843.
[38] X. Chen, K. Muthoosamy, A. Pfisterer, B. Neumann, T. Weil, *Bioconjugate Chem.* **2012**, *23*, 500–508.
[39] X. Chen, L. Henschke, Q. Wu, K. Muthoosamy, B. Neumann, T. Weil, *Org. Biomol. Chem.* **2013**, *11*, 353–361.
[40] T. Kortemme, T. E. Creighton, *J. Mol. Biol.* **1995**, *253*, 799–812.
[41] S. M. Marino, V. N. Gladyshev, *J. Biol. Chem.* **2012**, *287*, 4419–4425.
[42] H. R. Kilgore, R. T. Raines, *J. Am. Chem. Soc.* **2018**, *140*, 17606–17611.

Manuscript received: May 4, 2022
Accepted manuscript online: May 8, 2022
Version of record online: June 1, 2022