Article:
Yates, Nicholas D.J., Warnes, Matthew E., Breetveld, Reuben et al. (2023) Preparation and Application of an Inexpensive α-Formylglycine Building Block Compatible with Fmoc Solid-Phase Peptide Synthesis. Organic Letters. ISSN: 1523-7052

https://doi.org/10.1021/acs.orglett.2c04059

OL Organic Letters

pubs.acs.org/OrgLett Letter

Preparation and Application of an Inexpensive α-Formylglycine Building Block Compatible with Fmoc Solid-Phase Peptide Synthesis

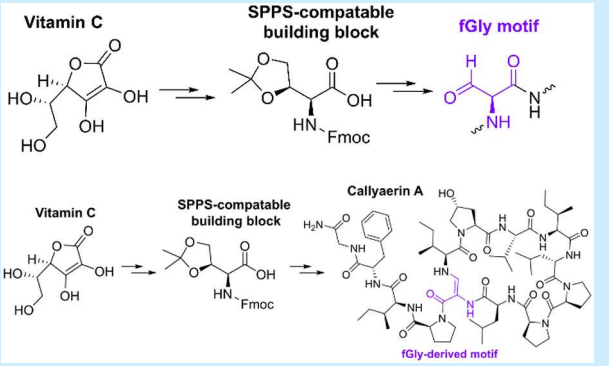
Published as part of the Organic Letters virtual special issue "Chemoselective Methods for Labeling and Modification of Peptides and Proteins".

Nicholas D. J. Yates,* Matthew E. Warnes, Reuben Breetveld, Christopher D. Spicer, Nathalie Signoret, and Martin Fascione*

Cite This: https://doi.org/10.1021/acs.orglett.2c04059 Read Online

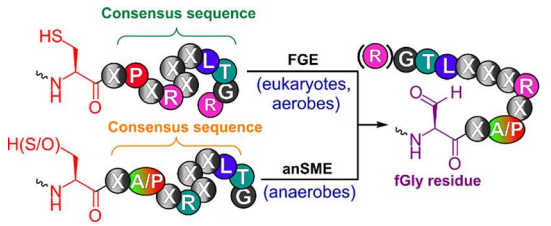
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: α-Formylglycine (fGly) is a rare residue located in the active site of sulfatases and serves as a precursor to pharmaceutically relevant motifs. The installation of fGly motifs into peptides is currently challenging due to degradation under the acidic and nucleophile-rich conditions accompanying resin cleavage during solid-phase peptide synthesis. We report the synthesis of acid- and nucleophile-tolerant α-formylglycine building blocks from vitamin C and use them to prepare callyaerin A, a macrocyclic peptide containing an fGly-derived motif.



In both eukaryotes and prokaryotes, enzymes capable of introducing α-formylglycine (fGly) residues (Scheme 1) into proteins prior to folding are essential for making post- or cotranslational modifications to type I sulfatases[1] and some phosphonate monoester hydrolases/phosphodiesterases.[2] A fGly residue is generated by the selective oxidation of a cysteine or serine residue embedded within a conserved consensus sequence (C/S)XPXRXXXLTG[1b,3] by either formylglycine generating enzymes (FGEs)[4] or anaerobic sulfatase maturating enzymes (anSMEs) (Scheme 1).[3b,5] Oxidation of this residue is used to introduce an active-site aldehyde functionality into the sulfatase/phosphonate monoester hydrolase/phosphodiesterases, which when hydrated to a geminal diol serves as a nucleophile, allowing these enzymes to cleave sulfate/phosphonate esters.[1,6] Incorporation of fGly consensus sequences into recombinant proteins can be used in conjugation with FGEs or anSMEs to produce fGly-labeled proteins, in which the fGly motif can serve as a unique bioconjugation handle.[7]

Scheme 1. Oxidation of Cysteine or Serine Residues by Either FGE or anSME Enzymes to Yield Formylglycine Residues[1b,3−5]

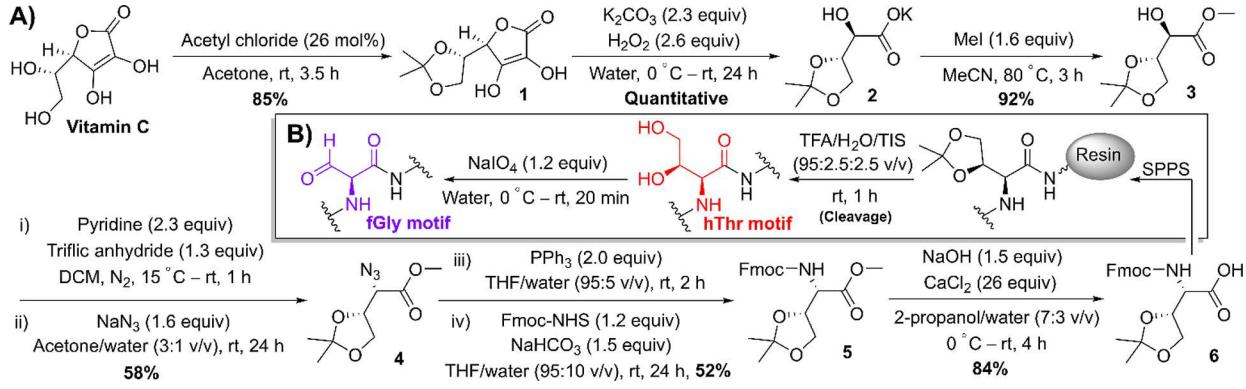


FGly-functionalized peptides are therefore useful in studies involving sulfatases, phosphonate monoester hydrolases/phosphodiesterases, FGEs, and an SMEs, and would also prove useful in the development of bioconjugation techniques intended for selectively targeting fGly residues.[7a,b,8] Additionally, fGly can also be used as a handle through which pharmaceutically important rigid diaminoacrylamide motifs can be installed into peptides.[9]

Whereas chemical approaches for fGly installation into peptides have been reported,[9a,b] these methods currently suffer from limitations associated with the intolerance of fGly to the acidic conditions and nucleophiles (e.g., silanes and thiol-based scavengers[9a] or cleaved 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) protecting groups) during the

Received: November 29, 2022

Scheme 2. (A) Preparation of Protected fGly Precursor 6 from Vitamin C. (B) Preparation of fGly-Functionalized Peptides Using Protected fGly Precursor 6



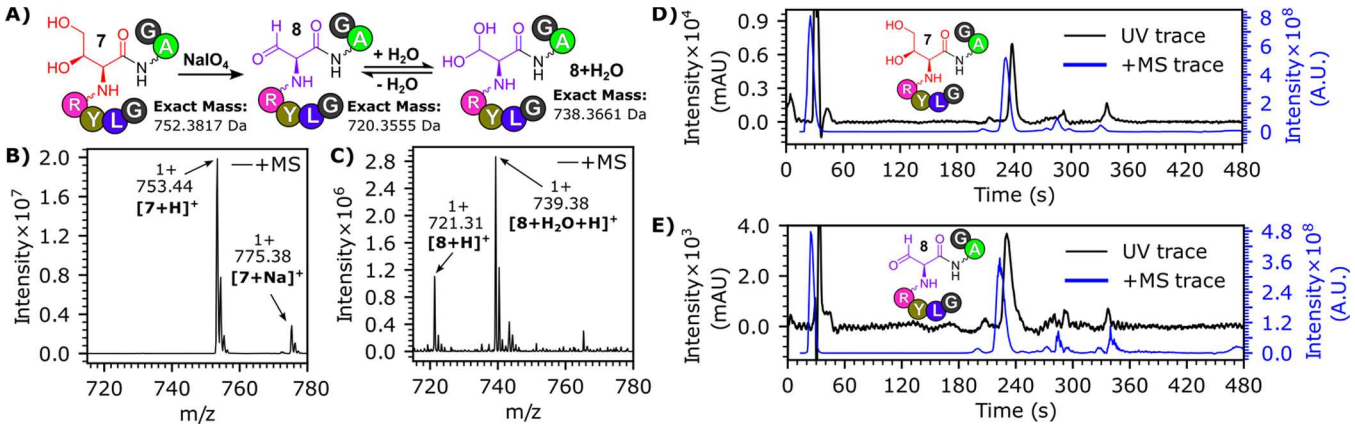


Figure 1. (A) Conversion of peptide **7** into peptide **8** using $NaIO_4$. (B) MS analysis of **7**. (C) MS analysis of **8**. (D) LC−MS chromatogram of **7**. (E) LC_-MS chromatogram of **8**. LC_-MS chromatograms show the BPC+All MS trace and the UV chromatogram shows the absorbance for wavelengths 210−400 nm. LC_MS was conducted using LC gradient A (see Figure S1).

resin-cleavage step of solid-phase peptide synthesis (SPPS) (see Figure S51).[9b,d] Herein we describe the design and total synthesis of an SPPS-compatible fGly-precursor building block from vitamin C: a cheap, readily available chiral feedstock (Scheme 2A). We demonstrate that this building block is compatible with conventional and microwave-assisted SPPS and prepare the pharmaceutically relevant peptide callyaerin A; a potent antituberculosis macrocyclic peptide which contains a (*Z*)-2,3-diaminoacrylamide moiety derived from an fGly residue inaccessible to FGE/anSME-based enzymatic efforts.[9d,10]

Oxidative cleavage of 1,2-diols[11] or 1,2-amino alcohols[12] using periodate is a mild and commonly used method for the installation of aldehydes into biomolecules. 1,2-Diols present in glycans and other biomolecules are highly susceptible to periodate oxidation,[11] yet are tolerant to the conditions required for resin cleavage after SPPS. We thus envisioned installing 1,2-diol-functionalized 4-hydroxy-L-threonine (hThr) residues into peptides as formyl glycine precursors for the subsequent generation of fGly residues after resin cleavage via mild sodium periodate oxidation (Scheme 2B). We hypothesized this approach would overcome issues associated with the intolerance of fGly residues to the conditions encountered during SPPS resin cleavage.

Methyl 3,4-*O*-isopropylidene-L-threonate **3** was identified as a commercially available chiral feedstock and was itself synthesized in-house in excellent yield using inexpensive reagents in three steps from vitamin C (Scheme 2A).[13] Subsequent substitution of the secondary alcohol in **3** via S_N2 attack with sodium azide, followed by Staudinger reduction, Fmoc-protection, and methyl ester hydrolysis, allowed *N*-Fmoc-protected SPPS-compatible hThr L-amino acid **6** to be accessed in a further four steps from **3**. The additional chiral center of the protected 1,2-diol motif allowed **6** to be verified as a pure L-amino acid by NMR.

fGly building block **6** was used to prepare a test peptide NH_2-Gly-Leu-Tyr-Arg-hThr-Ala-Gly-COOH **7** using Fmoc SPPS in a total yield of 95%, with HPLC analysis of **7** showing it to be of excellent purity (Figure 1). No acid-catalyzed breakdown products were observed during the cleavage of **7** from the peptide resin, and no side-products derived from the reaction of fGly with the cleaved Pbf group of the Arg residue were observed. Conversion of the hThr residue of **7** into fGly was then readily achieved via treatment of a 2 mM aqueous solution of **7** with 1.2 equiv of sodium metaperiodate ($NaIO_4$), yielding peptide **8** (Figure 1). Peptides **7** and **8** were found to have the same retention time in LC−MS (most likely due to the similar polarities of hydrated fGly and hThr motifs), but mass spectrometry peaks attributable to **7** were fully lost within 20 min of treatment with $NaIO_4$, leaving only signals attributable to **8** (Figure 1). Additionally, while oxidation of serine/threonine residues to fGly using RuO_4 is known to lead to side-products via oxidative scission/retroaminal fragmentation,[9a,14] no such

https://doi.org/10.1021/acs.orglett.2c04059
Org. Lett. XXXX, XXX, XXX−XXX

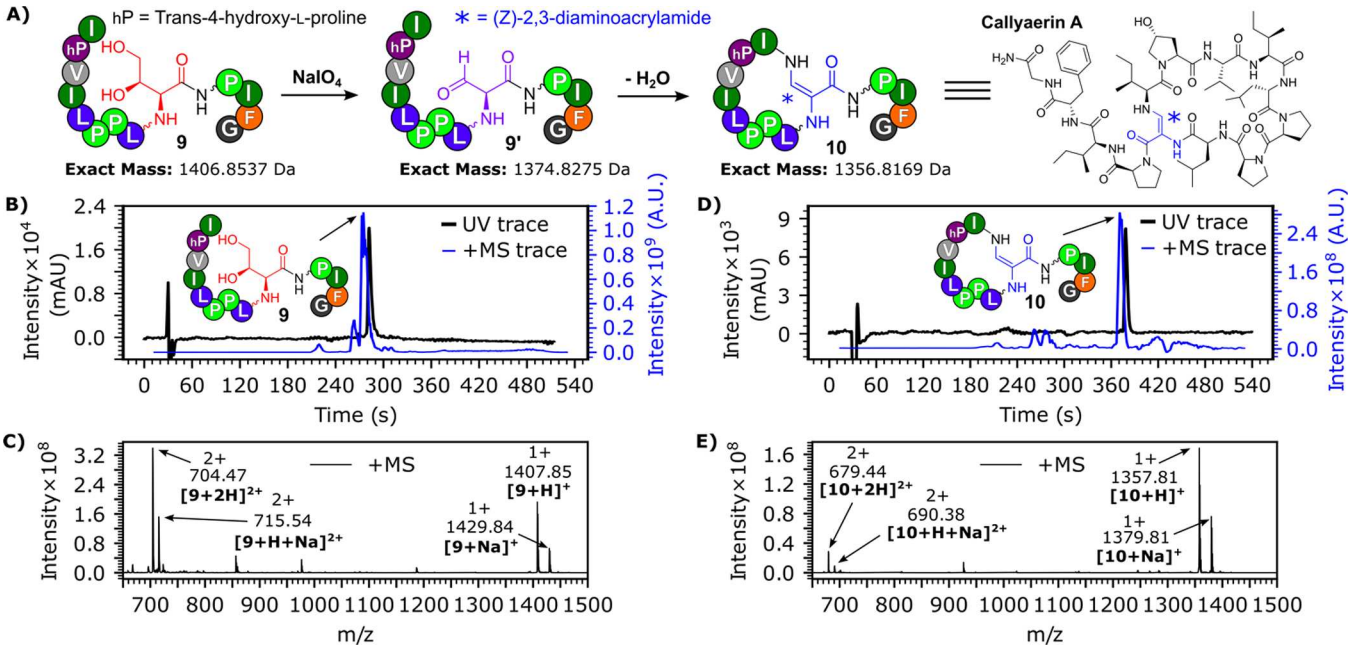


Figure 2. (A) Conversion of linear callyaerin A precursor **9** into callyaerin A **10** (via **9′**) using $NaIO_4$. (B) LC−MS chromatogram of **9**. (C) Mass spectrum of **9**. (D) LC−MS chromatogram of **10**. (E) Mass spectrum of **10**. Note that the peak at m/z 926.57 corresponds to a b_2 fragment ion of **10** (Figure S50). LC−MS chromatograms show the BPC+All MS trace and the UV chromatogram shows the absorbance for wavelengths 210−400 nm. LC−MS was conducted using LC gradient B (see Figure S2). For MS evidence of **9′** see Figure S53.

byproducts were detected after the treatment of **7** with $NaIO_4$ (see Figure S52).

Having verified the compatibility of **6** with SPPS, and with conditions established for fGly installation via periodate treatment, we applied the methodology to the synthesis of callyaerin A **10**,[9c,d] which contains a rare (*Z*)-2,3-diaminoacrylamide motif derived from an fGly residue (Figure 2).[9a,d] The only previous synthesis of **10** used an alternative Fmoc-protected building block but reported problems associated with the epimerization of this building block in SPPS, in addition to acid-catalyzed fragmentation of the fGly-functionalized linear callyaerin A peptide upon resin cleavage.[9d] By contrast, we hypothesized that a comparable linear peptide constructed using **6** would be far less prone to acid-catalyzed fragmentation during resin cleavage.

Microwave-assisted fully automated Fmoc SPPS was used to prepare linear callyaerin A precursor **9** (Figure 2), verifying that **6** is suitable for use in microwave-assisted peptide synthesis. A 2 mM solution of **9** was then treated with 1.2 equiv of $NaIO_4$, and the hThr residue of **9** was cleanly converted to fGly in a one-pot reaction within 20 min, yielding a mixture of linear callyaerin A **9′** and callyaerin A **10** (see Figure S53). Complete cyclization of **9′** into **10** was facilitated via the extraction of the peptide mixture into MeCN and the addition of 0.1% formic acid and $MgSO_4$. Due to the profound change in the shape accompanying cyclization, **9** and **10** have distinct retention times as well as distant mass spectrometry peaks (Figure 2).

In summary, Fmoc-protected hThr building block **6** can be readily synthesized from vitamin C using inexpensive reagents and used in SPPS to install hThr residues into peptides. Utilization of hThr as a precursor to fGly can circumvent incompatibility issues associated with the intolerance of fGly toward the conditions/species encountered during SPPS, and the hThr residue can then be cleanly converted to fGly off-resin using $NaIO_4$. We demonstrated that our approach can not only be used to prepare fGly-functionalized peptides but also used to prepare cyclic peptides bearing pharmaceutically-important rigidifying motifs. Although under harsh conditions amino acids such as Met, Trp, Tyr, and Ser/Thr can be susceptible periodate oxidation,[15] by performing reactions at neutral pH and low temperature with controlled reaction stoichiometry, akin to the conditions used here, these residues have been shown to tolerate $NaIO_4$.[11a,16] Furthermore, addition of excess Met to the reaction has been shown to further reduce overoxidation of peptides with no loss of biological activity, notably in the case of the polypeptide chemokine RANTES which contains Met and also multiple disulfide forming Cys residues,[17] which are highly susceptible to periodate oxidation.[18] We therefore propose our building block may have broad applications in the fields of bioconjugation and pharmacology, particularly as bioconjugations using formylglycine are currently underexplored, the biological mechanisms of formylglycine installation has not yet been unequivocally elucidated[19] and formylglycine-derived motifs are present in many pharmaceutically-relevant compounds, including the tuberactinomycin[9a,20] and callyaerin[9a,c,d,10,21] families of antibiotics.

■ ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are available in the published article and its Supporting Information.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.orglett.2c04059.

Experimental procedures and spectroscopic and analytical data for all new compounds, including copies of the NMR spectra (PDF)

https://doi.org/10.1021/acs.orglett.2c04059
Org. Lett. XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Authors

Nicholas D. J. Yates − *Department of Chemistry, University of York, York YO10 5DD, U.K.*; orcid.org/0000-0002-4871-2133; Email: nicholas.yates@york.ac.uk

Martin Fascione − *Department of Chemistry, University of York, York YO10 5DD, U.K.*; orcid.org/0000-0002-0066-4419; Email: martin.fascione@york.ac.uk

Authors

Matthew E. Warnes − *Department of Chemistry, University of York, York YO10 5DD, U.K.*

Reuben Breetveld − *Department of Chemistry, University of York, York YO10 5DD, U.K.*

Christopher D. Spicer − *Department of Chemistry, University of York, York YO10 5DD, U.K.*; orcid.org/0000-0001-8787-578X

Nathalie Signoret − *Hull York Medical School, University of York, York YO10 5DD, U.K.*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.orglett.2c04059

Author Contributions

All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Ed Bergstrom and The York Centre of Excellence in Mass Spectrometry. The York Centre of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1). This work was supported by The University of York, the EPSRC (summer vacation bursary for M.E.W. and EP/V044303/1) and UKRI (EP/X023680/1).

REFERENCES

(1) (a) Carlson, B. L.; Ballister, E. R.; Skordalakes, E.; King, D. S.; Breidenbach, M. A.; Gilmore, S. A.; Berger, J. M.; Bertozzi, C. R. *J. Biol. Chem.* **2008**, *283*, 20117−20125. (b) Appel, M. J.; Bertozzi, C. R. *ACS Chem. Biol.* **2015**, *10*, 72−84.

(2) Knop, M.; Engi, P.; Lemnaru, R.; Seebeck, F. P. *ChemBioChem.* **2015**, *16*, 2147−2150.

(3) (a) Kruger, T.; Weiland, S.; Boschanski, M.; Sinha, P. K.; Falck, G.; Muller, K. M.; Dierks, T.; Sewald, N. *ChemBioChem.* **2019**, *20*, 2074−2078. (b) Grove, T. L.; Lee, K.-H.; Clair, J. St.; Krebs, C.; Booker, S. J. *Biochemistry* **2008**, *47*, 7523−7538.

(4) Appel, M. J.; Meier, K. K.; Lafrance-Vanasse, J.; Lim, H.; Tsai, C.-L.; Hedman, B.; Hodgson, K. O.; Tainer, J. A.; Solomon, E. I.; Bertozzi, C. R. *Proc. Nat. Acad. Sci.* **2019**, *116*, 5370.

(5) (a) Goldman, P. J.; Grove, T. L.; Sites, L. A.; McLaughlin, M. I.; Booker, S. J.; Drennan, C. L. *Proc. Nat. Acad. Sci.* **2013**, *110*, 8519. (b) Benjdia, A.; Subramanian, S.; Leprince, J.; Vaudry, H.; Johnson, M. K.; Berteau, O. *FEBS J.* **2010**, *277*, 1906−1920.

(6) (a) Jonas, S.; van Loo, B.; Hyvönen, M.; Hollfelder, F. *J. Mol. Biol.* **2008**, *384*, 120−136. (b) Carlson, B. L.; Ballister, E. R.; Skordalakes, E.; King, D. S.; Breidenbach, M. A.; Gilmore, S. A.; Berger, J. M.; Bertozzi, C. R. *J. Biol. Chem.* **2008**, *283*, 20117−20125.

(7) (a) Appel, M. J.; Bertozzi, C. R. *ACS Chem. Biol.* **2015**, *10*, 72−84. (b) Krüger, T.; Dierks, T.; Sewald, N. *Biol. Chem.* **2019**, *400*, 289−297. (c) York, D.; Baker, J.; Holder, P. G.; Jones, L. C.; Drake, P. M.; Barfield, R. M.; Bleck, G. T.; Rabuka, D. *BMC Biotechnology* **2016**, *16*, 23. (d) Krüger, T.; Weiland, S.; Falck, G.; Gerlach, M.; Boschanski, M.; Alam, S.; Müller, K. M.; Dierks, T.; Sewald, N. *Angew. Chem., Int. Ed.* **2018**, *57*, 7245−7249.

(8) Tiemann, M.; Nawrotzky, E.; Schmieder, P.; Wehrhan, L.; Bergemann, S.; Martos, V.; Song, W.; Arkona, C.; Keller, B. G.; Rademann, J. *Chem.—Eur. J.* **2022**, *28*, e202201282.

(9) (a) Zhang, S.; Rodriguez, L. M. D. L.; Harris, P. W. R.; Brimble, M. A. *Asian J. Org. Chem.* **2017**, *6*, 1180−1190. (b) Rush, J.; Bertozzi, C. R. *Org. Lett.* **2006**, *8*, 131−134. (c) Ibrahim, S. R. M.; Min, C. C.; Teuscher, F.; Ebel, R.; Kakoschke, C.; Lin, W.; Wray, V.; Edrada-Ebel, R.; Proksch, P. *Biorg. Med. Chem.* **2010**, *18*, 4947−4956. (d) Zhang, S.; De Leon Rodriguez, L. M.; Leung, I. K. H.; Cook, G. M.; Harris, P. W. R.; Brimble, M. A. *Angew. Chem., Int. Ed.* **2018**, *57*, 3631−3635.

(10) Daletos, G.; Kalscheuer, R.; Koliwer-Brandl, H.; Hartmann, R.; de Voogd, N. J.; Wray, V.; Lin, W.; Proksch, P. *J. Nat. Prod.* **2015**, *78*, 1910−1925.

(11) (a) Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992**, *3*, 138−146. (b) Mikolajczyk, S. D.; Meyer, D. L.; Starling, J. J.; Law, K. L.; Rose, K.; Dufour, B.; Offord, R. E. *Bioconjugate Chem.* **1994**, *5*, 636−646. (c) Clamp, J. R.; Hough, L. *Biochem. J.* **1965**, *94*, 17−24. (d) Stanley, P.; Schachter, H.; Taniguchi, N. In *Essentials of Glycobiology*; Varki, A., Cummings, R., Esko, J., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009. (e) Hounsell, E. F.; Davies, M. J.; Smith, K. D. In *The Protein Protocols Handbook*; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 2009; pp 1215−1217;. (f) Kirkeby, S. *Biotechnic & Histochemistry* **2016**, *91*, 1−8. (g) Norgard, K. E.; Han, H.; Powell, L.; Kriegler, M.; Varki, A.; Varki, N. M. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 1068−1072. (h) Yates, N. D.; Dowsett, M. R.; Bentley, P.; Dickenson-Fogg, J. A.; Pratt, A.; Blanford, C. F.; Fascione, M. A.; Parkin, A. *Langmuir* **2020**, *36*, 5654−5664. (i) Reuter, G.; Schauer, R.; Szeiki, C.; Kamerling, J. P.; Vliegenthart, J. F. G. *Glycoconjugate J.* **1989**, *6*, 35−44. (j) Ramya, T. N. C.; Weerapana, E.; Cravatt, B. F.; Paulson, J. C. *Glycobiology* **2013**, *23*, 211−221. (k) Thaysen-Andersen, M.; Larsen, M. R.; Packer, N. H.; Palmisano, G. *RSC Adv.* **2013**, *3*, 22683−22705.

(12) (a) Yates, N. D. J.; Akkad, S.; Noble, A.; Keenan, T.; Hatton, N. E.; Signoret, N.; Fascione, M. A. *Green Chem.* **2022**, *24*, 8046−8053. (b) Spears, R. J.; Brabham, R. L.; Budhadev, D.; Keenan, T.; McKenna, S.; Walton, J.; Brannigan, J. A.; Brzozowski, A. M.; Wilkinson, A. J.; Plevin, M.; Fascione, M. A. *Chem. Sci.* **2018**, *9*, 5585−5593. (c) Spears, R. J.; Fascione, M. A. *Org. Biomol. Chem.* **2016**, *14*, 7622−7638. (d) Keenan, T.; Spears, R. J.; Akkad, S.; Mahon, C. S.; Hatton, N. E.; Walton, J.; Noble, A.; Yates, N. D.; Baumann, C. G.; Parkin, A.; Signoret, N.; Fascione, M. A. *ACS Chem. Biol.* **2021**, *16*, 2387−2400. (e) Brabham, R. L.; Keenan, T.; Husken, A.; Bilsborrow, J.; McBerney, R.; Kumar, V.; Turnbull, W. B.; Fascione, M. A. *Org. Biomol. Chem.* **2020**, *18*, 4000−4003.

(13) Kou, Q.; Wang, T.; Zou, F.; Zhang, S.; Chen, Q.; Yang, Y. *Eur. J. Med. Chem.* **2018**, *151*, 98−109.

(14) Ranganathan, D.; Vaish, N. K.; Shah, K. *J. Am. Chem. Soc.* **1994**, *116*, 6545−6557.

(15) El-Mahdi, O.; Melnyk, O. *Bioconjugate Chem.* **2013**, *24* (5), 735−765.

(16) Dixon, H. B.; Weitkamp, L. R. *Biochem. J.* **1962**, *84*, 462−468.

(17) Simmons, G.; Clapham, P. R.; Picard, L.; Offord, R. E.; Rosenkilde, M. M.; Schwartz, T. W.; Buser, R.; Wells, T. N. C.; Proudfoot, A. E. I. *Science* **1997**, *276* (5310), 276−279.

(18) Liu, B.; Burdine, L.; Kodadek, T. *J. Am. Chem. Soc.* **2006**, *128*, 15228−15235.

(19) Ennemann, E. C.; Radhakrishnan, K.; Mariappan, M.; Wachs, M.; Pringle, T. H.; Schmidt, B.; Dierks, T. *J. Biol. Chem.* **2013**, *288*, 5828−5839.

(20) (a) DeMong, D. E.; Williams, R. M. *J. Am. Chem. Soc.* **2003**, *125*, 8561−8565. (b) Stanley, R. E.; Blaha, G.; Grodzicki, R. L.; Strickler, M. D.; Steitz, T. A. *Nat. Struct. Mol. Biol.* **2010**, *17*, 289−293.
(21) Yates, N. D. J.; Warnes, M.; Breetveld, R.; Spicer, C.; Signoret, N.; Fascione, M. A. *ChemRxiv.* **2022**, DOI: 10.26434/chemrxiv-2022-z84gd.

https://doi.org/10.1021/acs.orglett.2c04059
Org. Lett. XXXX, XXX, XXX−XXX