

Peptides

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 14319–14323

International Edition: doi.org/10.1002/anie.202100818

German Edition: doi.org/10.1002/ange.202100818

Engineering of a Peptide α -N-Methyltransferase to Methylate Non-Proteinogenic Amino Acids

Haigang Song, Antony J. Burton, Sally L. Shirran, Jūratė Fahrīg-Kamaraukaitė, Hannelore Kaspar, Tom W. Muir, Markus Künzler, and James H. Naismith*

Abstract: Introduction of α -N-methylated non-proteinogenic amino acids into peptides can improve their biological activities, membrane permeability and proteolytic stability. This is commonly achieved, in nature and in the lab, by assembling pre-methylated amino acids. The more appealing route of methylating amide bonds is challenging. Biology has evolved an α -N-automethylating enzyme, OphMA, which acts on the amide bonds of peptides fused to its C-terminus. Due to the ribosomal biosynthesis of its substrate, the activity of this enzyme towards peptides with non-proteinogenic amino acids has not been addressed. An engineered OphMA, intein-mediated protein ligation and solid-phase peptide synthesis have allowed us to demonstrate the methylation of amide bonds in the context of non-natural amides. This approach may have application in the biotechnological production of therapeutic peptides.

Peptide cyclization, backbone N-methylation, and the use of D and non-proteinogenic amino acids are common strategies in medicinal chemistry to enhance the metabolic stability, cell permeability, and bioavailability of synthetic peptides.^[1–6] Methylation is a particularly desirable modification since it reduces the polarity of the amide bond, introduces conformational constraints and sequential introduction appears additive in terms of increasing membrane permeability.^[6–8] Cyclosporine A, the widely prescribed immunosuppressant, exemplifies this with its seven backbone N-methylations, an aminobutyric acid (Abu), N-methylated butenyl-threonine (Bmt) and D-alanine residues.^[9] Cyclosporin A is a non-ribosomal peptide natural product (NRPs) where methylation is carried out on the building blocks by methyltransferase domains prior to peptide bond formation.^[10,11] Manipulating NRP synthetases to generate novel N-methylated unnatural

amino acids is challenging due to the complexity of the assembly lines.^[12,13] N-methylated amino acids have been incorporated into peptides by expanding or reprogramming the genetic code and by using pre-charged N-methyl amino-acyl-tRNAs in cell-free ribosomal translation systems.^[14–18] Traditionally this approach had low efficiency but the novel flexizyme system has been able to produce polymethylated peptides using tRNA loaded with methylated amino acid building blocks.^[17–22] These approaches do however not methylate the peptide bond itself and, thus, do not utilize the full diversity available through chemical peptide synthesis. Enzymatic methylation of the amide bond has recently been discovered^[23,24] (Figure 1 a), but its potential for processing amide bonds involving non-proteinogenic amino acids has not been explored. The enzymatic methylation of chemically synthesized peptides creates the possibility of highly diverse molecules.

Structural analysis of OphMA revealed that the substrate peptide is mainly recognized via hydrogen bonds to substrate backbone carbonyls^[26–28] with side chains of substrate playing a minor role. OphMA has been shown to prefer hydrophobic residues (Gly, Ala, Leu, Ile), small hydrophilic residues (Thr, Ser), tolerate Phe and (albeit inefficiently) methylate Glu.^[23,24,29] OphMA homologs from other fungi are able to methylate different length core peptides and process His, Cys, Tyr, Gln, and Asp residues.^[28,30] Since no *in trans* activity of OphMA has been identified,^[27,31] the substrate scope of these methyltransferases is limited to the 20 proteinogenic amino acids. Genetic code engineering could introduce non-natural amino acids into OphMA core peptide but limits diversity.^[32–35] We therefore chose to develop a system that combined peptide synthesis with enzymatic amide methylation. We selected expressed protein ligation (EPL)^[36,37] to

[*] Dr. H. Song, Prof. J. H. Naismith
Division of Structural Biology
Wellcome Centre for Human Genetics
Roosevelt Drive, Oxford OX3 7BN (UK)
and
The Research Complex at Harwell
Harwell Campus, Oxford OX11 0FA (UK)
and
The Rosalind Franklin Institute
Harwell Campus, Oxford OX11 0FA (UK)
E-mail: naismith@strubi.ox.ac.uk
Dr. A. J. Burton, Prof. T. W. Muir
Department of Chemistry, Frick Chemistry Laboratory
Princeton University
Princeton, NJ (USA)

Dr. S. L. Shirran
Biomedical Sciences Research Complex, North Haugh
University of St. Andrews
Fife KY16 9ST (UK)

Dr. J. Fahrīg-Kamaraukaitė, Dr. H. Kaspar, Prof. M. Künzler
Department of Biology, Institute of Microbiology
Eidgenössische Technische Hochschule (ETH) Zürich
Zürich (Switzerland)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/anie.202100818>.

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

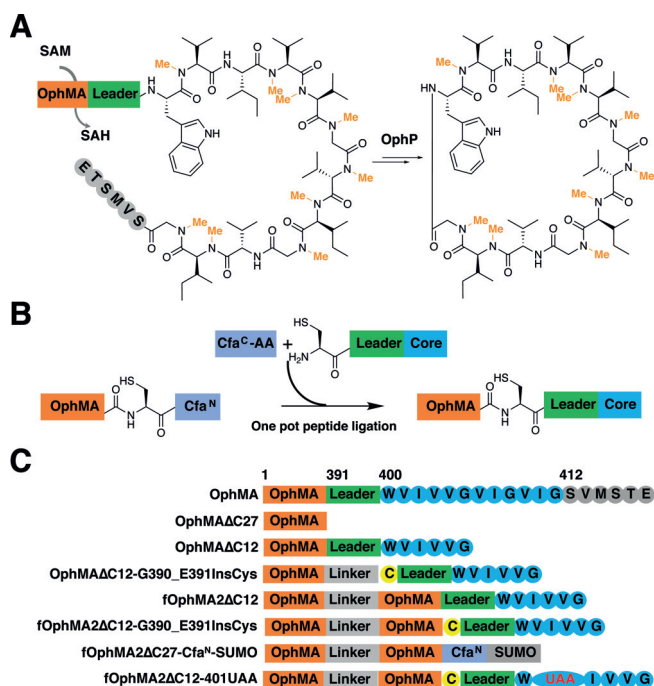


Figure 1. OphMA- and split-intein-mediated peptide ligation.

A) OphMA catalyzes nine successive backbone N-methylations of its C-terminal peptide which is converted to omphalotin A by OphP (macrocyclase).^[23–25] The leader peptide is shown green, the follower as gray circles. B) The scheme for one-pot core peptide (blue) ligation to OphMA mediated by split intein Cfa^N and mutated Cfa^C (denoted Cfa^C-AA) inserts a cysteine at the ligation site (see also Figure S1). C) Key to constructs. Native OphMA (numbers show the first residue of each region); OphMAΔC27 and OphMAΔC12 (27 or 12 C-terminal residues deleted); OphMAΔC12-G390_E391InsCys as OphMAΔC12 but cysteine (yellow) inserted between 390 and 391; fOphMA2ΔC12 is the fused dimer with 12 C-terminal residues deleted; fOphMA2ΔC12-G390_E391InsCys is fOphMA2ΔC12 with cysteine insertion; fOphMA2ΔC27-Cfa^N-SUMO intein and sumo added to C-terminus; fOphMA2ΔC12-401UAA the post ligation protein.

couple a synthetic peptide to the enzyme using the Cfa split intein system (consensus DnaE intein sequence) as it has excellent thermal and chaotropic stability (Figure 1b, Figure S1).^[37–43]

OphMA functions as an interlocked dimer forming a catenane arrangement with C-terminal substrate of one monomer folded into the active tunnel of the other monomer.^[26] The catenane arrangement of OphMA seems certain to be functionally required. Guided by this structural insight, we truncated OphMA after Gly390 in order to preserve the last α -helix (S380–E388, thus predicted structural stability), yielding OphMAΔC27 to which the Cfa intein would be added (Figure 1c). Ligation of peptides to this system would result in a cysteine insertion between residues 390 and 391 of OphMAΔC12 (Figure 1c). Construction of fused dimers has been shown to improve the expression level, stability and activity of some enzymes.^[44–46] Again influenced by the structure, we constructed a fused OphMA dimer (fOphMA2) (Figure 1c). The fused dimer consists of Met1 to Asn377 of one monomer linked by (GGGGS)₆ to the second (variable) monomer.

The C-terminal peptide of OphMA is insoluble, therefore a soluble test sequence **WVIVVG** was chosen for the core peptide. OphMAΔC12, OphMAΔC12-G390_E391InsCys and fOphMA2ΔC12-G390_E391InsCys (Figure 1c) were expressed in *E. coli* to evaluate enzyme activity on the test sequence made during ribosomal synthesis.^[26,29] OphMAΔC12 was purified as a mixture with zero, one and two methylations (Figure 2a) confirming it was active. Cysteine insertions showed improved methylation when compared to OphMAΔC12 with more mono- and di-methylation. The fused dimer was purified as a well-folded methylated protein (Figure S2) with improved thermal stability (Figure S3). Both intact mass and MS² analysis identified that Val401 and Val403 were methylated (expected from previous work (Figure S7 & S8)). All three constructs showed the expected methylation pattern for this test sequence; the two C-terminal residues are methylated.^[23,24,26–29] Since the engineered enzymes were active, we proceeded to ligate a peptide containing the test sequence (CEEASQNGFP-**WVIVVG**, peptide 1, Table S1).

Split-intein fusion variants were constructed with a C-terminal SUMO fusion tag (simplifying monitoring) for the monomer (OphMAΔC27-Cfa^N-SUMO) and fused dimer (fOphMA2ΔC27-Cfa^N-SUMO, Figure 1c). Both proteins were purified from *E. coli*^[26,29] and showed 60% ligation efficiency in 4 h and reached around 80% within 20 h (estimated by SDS-PAGE gels, Figure S4).

ESI-TOF MS determined the molecular mass as 44 514 Da for OphMAΔC12-401 Val (a peak at 44 529 Da corresponds to oxidation) and 89 557 Da for fOphMA2ΔC12-401 Val, the expected masses for ligated proteins (Figure 2b, S6). Peaks at 42 900 Da in OphMA and 87 944 Da in fOphMA2 corresponded to unligated product, likely a result of thioester hydrolysis or cleavage. Both OphMAΔC12-401 Val and fOphMA2ΔC12-401 Val were incubated with SAM for three days as previously described^[26,30] and analysis of Glu-C-digested C-terminal peptides returned twenty percent of the mono-methylated species (Figure 2c, S7). MS² confirmed that Val401 was methylated in both proteins, and the di-methylated species was also detected for fOphMA2 (Figure 2c, S8) suggesting it was more active. The *in vitro* activities of the ligated variants were lower than their translated counterparts but validated the approach. Since the fused protein was more stable, readily ligated and more active than monomer we pursued it.

A panel of synthetic peptides with a range of non-proteinogenic amino acids at position 401 were ligated to fOphMA2ΔC27-Cfa^N-SUMO and the extent of methylation both at the unnatural amino acid and elsewhere in the core peptide judged by mass spectrometry are reported in Table S1, Figure 3 and S6, S7, S8. Interestingly, both Nva and Nle were more efficiently methylated (by around 4-fold) than valine. The bioorthogonal reactive alkyne containing Pra was also well tolerated. MS² analysis confirmed that in addition to methylation at Val401Pra there was additional methylation on Ile402 or Val403 (Figure S7, S8). Introduction of larger non-aromatic cyclic hydrophobic side chains reduced methylation with 6% mono-methylation at Val401Cha, 11% mono-methylation at Val401Chg and 21% mono-methylation

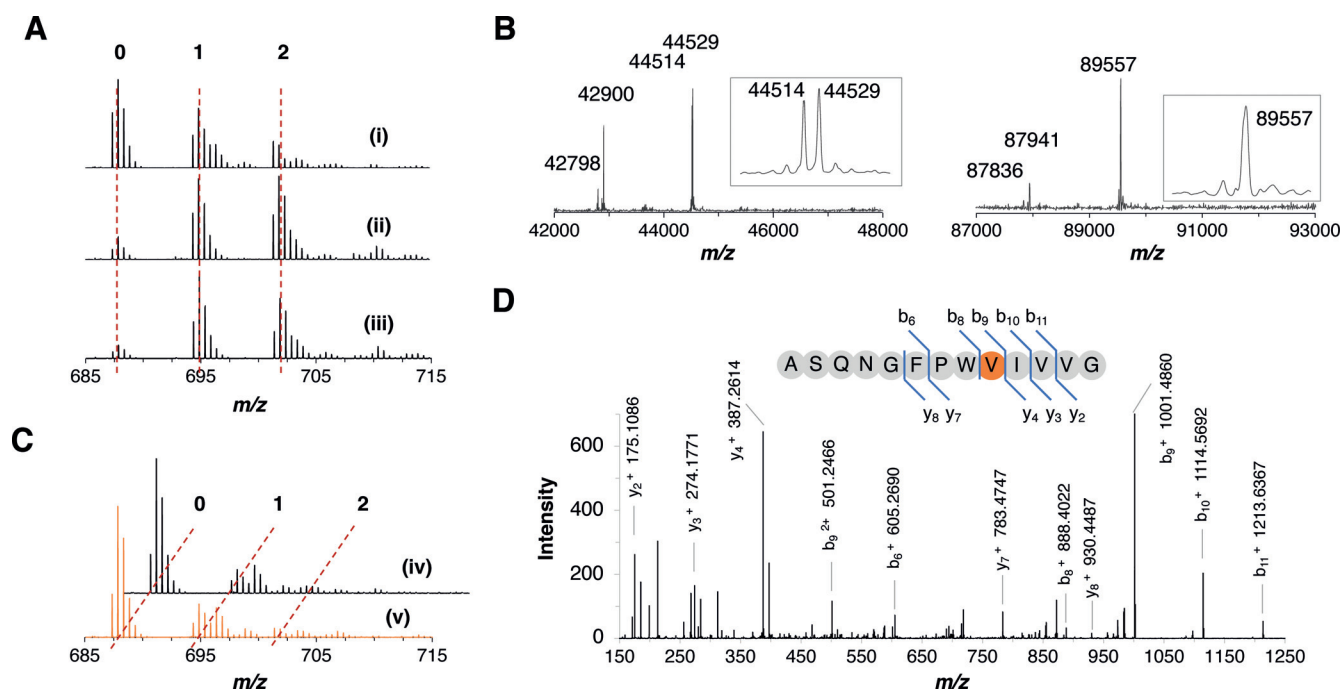


Figure 2. OphMA catalyzes backbone N-methylation of ligated synthetic peptides. A) MS analysis of C-terminal peptide of the expressed OphMA Δ C12 (i), OphMA Δ C12-G390_E391insCys (ii) and fOphMA2 Δ C12-G390_E391insCys (iii). B) Intact mass analysis of ligated products, OphMA Δ C12-401Val (left) and fOphMA2 Δ C12-401Val (right). C) MS analysis of C-terminal peptide confirmed N-methylation for OphMA Δ C12-401Val (iv) and fOphMA2 Δ C12-401Val (v). Both are ligated protein with cysteine insertion between Gly390 and Glu391, and incubated with SAM for three days. D) MS² spectrum confirms the N-methylated site of fOphMA2 Δ C12-401Val.

at Val401Cpg (Figure 3, Table S1, Figure S7, S8). OphMA has previously been shown to methylate phenylalanine in different positions albeit with reduced efficiency.^[29] MS and MS² spectra showed that fOphMA2 methylated various aromatic amino acids ranging from 7% methylation of 4-F-Phe to 40% with 2-Pal (Table S1, Figure 3 and S7, S8). Detectable methylation at Val401Phg was also seen (Table S1 and Figure S7, S8). Very large side chains (4-Cl-Phe, 4-I-Phe, 4-NO₂-Phe) prevented methylation. Homoserine was well tolerated and allowed a small amount of additional methylation (Ile402 or Val403, Table S1, Figure S7, S8). The highly branched penicillamine side chain was comparably methylated to valine (Table S1, Figure 3 and S7, S8). The positively charged diaminobutyric acid (Dab) yielded mono-methylated protein, a surprise since arginine blocks methylation in native OphMA. MS² showed two mono-methylated species, one with methylated Dab and the other with methylated Ile402 (Figure S8). Residues Aib, D-Ala, D-Val and D-Thr all completely abolished methylation. A variety of non- α amino acids were tested with only small amounts of mono-methylation observed for protein with Ahx or Gaba but MS² suggested that it was Ile402 or Val403 which were methylated (Figure S7, S8). We attribute the lack of ligation product with peptide **30** (CEEASQNGFP-**WVIVGVIG**) to the peptide's insolubility (Figure S6).

Since we have ligated the peptide to an already folded protein, the residues in the peptide must enter the active site. This was a surprising finding since it requires the enzyme to

undergo a large conformational change. The structure had implied substrate peptide entered the active site during the folding.^[29] In fact, the enzyme must undergo a conformational change in the active site. The ability to undergo this change without unfolding may explain why a fused dimeric enzyme behaved better. Our data showed the enzyme fOphMA2 was able to process hydrophobic side chains including aromatics but with a preference for smaller unbranched side chains. Despite the mismatch in polarity, we observed the methylation of a small positively charged residue, suggesting small size followed by hydrophobicity dominate substrate tolerance. Structural data had suggested D-configured amino acids would clash at the active site and indeed none were tolerated. The inefficiency of methylation of Phg suggested that backbone flexibility is important, the first experimental evidence to support the flip and translate model for OphMA hypermethylation.^[26,28] The lack of activity on non- α amino acids is likely due to disruption of the critical hydrogen bonding interactions between the substrate peptide and enzyme.^[26–28]

This is the first report of enzymatic N-methylation of amide bonds involving non-natural amino acids. Moreover, the fused protein appears to be a useful tool for further exploration, since it was able to process substrates that are added after the enzyme has folded. The fused enzyme is more stable and appears to be more active. Refinement of this strategy could provide access to entirely novel multiply backbone N-methylated peptides.

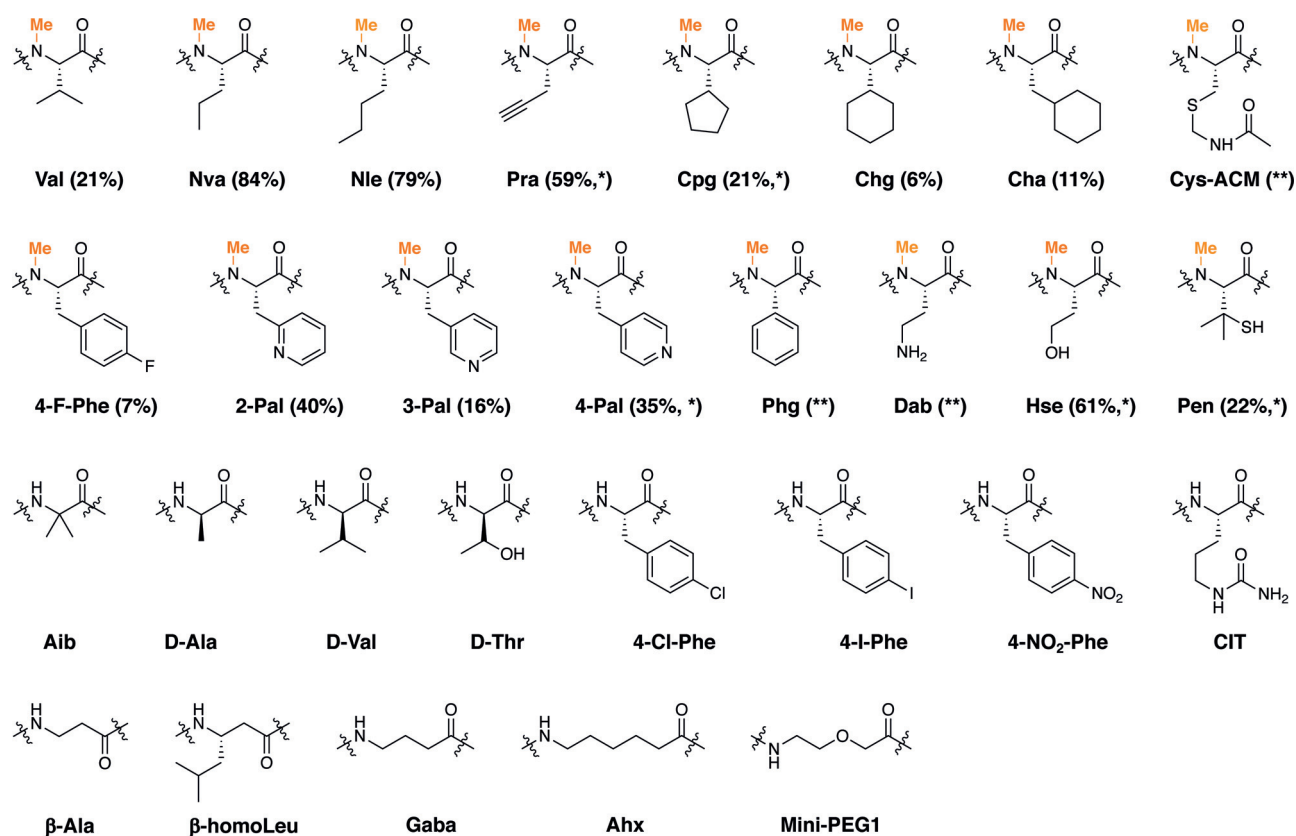


Figure 3. Non-proteinogenic amino acids N-methylated by OphMA. The methylation was determined by MS² of the Glu-C digested fOphMA2-ligated protein. Values in the bracket represent relative percentage of mono-methylation species with non-proteinogenic site methylated and are calculated by integrating all methylated species observed. * Di-methylation was also detected; ** indicates methylation of the residue but low abundance.

Acknowledgements

We thank Dr. Andrew Quigley and Nadisha Gamage of Membrane Protein lab for the help with thermal shift (202892/Z/16/Z), Wael Kamel, Prof. Shabaz Mohammed, Dr. Mateusz Imiolek for help with MS analysis. This work was supported by grants to M.K. (CTI/Innosuisse 25951.2) and J.H.N. (BBSRC BB/R018189/1 and ERC 339367). A.J.B. is a Damon Runyon Fellow of the Damon Runyon Cancer Research Foundation (DRG-2283-17).

Conflict of interest

M. Künzler is co-inventor of patent WO2017174760A1.

Keywords: cyclic peptide · non-proteinogenic amino acids · RiPPs · split intein · α -N-methylation

- [1] A. T. Bockus, J. A. Schwochert, C. R. Pye, C. E. Townsend, V. Sok, M. A. Bednarek, R. S. Lokey, *J. Med. Chem.* **2015**, *58*, 7409–7418.
- [2] J. L. Hickey, S. Zaretsky, M. A. St Denis, S. K. Chakka, M. M. Morshed, C. C. G. Scully, A. L. Roughton, A. K. Yudin, *J. Med. Chem.* **2016**, *59*, 5368–5376.
- [3] A. Furukawa, C. E. Townsend, J. Schwochert, C. R. Pye, M. A. Bednarek, R. S. Lokey, *J. Med. Chem.* **2016**, *59*, 9503–9512.

- [4] R. Jwad, D. Weissberger, L. Hunter, *Chem. Rev.* **2020**, *120*, 9743–9789.
- [5] T. Rezai, B. Yu, G. L. Millhauser, M. P. Jacobson, R. S. Lokey, *J. Am. Chem. Soc.* **2006**, *128*, 2510–2511.
- [6] J. Chatterjee, C. Gilon, A. Hoffman, H. Kessler, *Acc. Chem. Res.* **2008**, *41*, 1331–1342.
- [7] J. Chatterjee, F. Rechenmacher, H. Kessler, *Angew. Chem. Int. Ed.* **2013**, *52*, 254–269; *Angew. Chem.* **2013**, *125*, 268–283.
- [8] A. F. B. Räder, F. Reichart, M. Weinmüller, H. Kessler, *Bioorg. Med. Chem.* **2018**, *26*, 2766–2773.
- [9] M. Dreyfuss, E. Härrä, H. Hofmann, H. Kobel, W. Pache, H. Tschertter, *Eur. J. Appl. Microbiol.* **1976**, *3*, 125–133.
- [10] A. Lawen, R. Zocher, *J. Biol. Chem.* **1990**, *265*, 11355–11360.
- [11] R. D. Süßmuth, A. Mainz, *Angew. Chem. Int. Ed.* **2017**, *56*, 3770–3821; *Angew. Chem.* **2017**, *129*, 3824–3878.
- [12] R. H. Baltz, *Nat. Biotechnol.* **2006**, *24*, 1533–1540.
- [13] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468–3496.
- [14] C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G. Schultz, *Science* **1989**, *244*, 182–188.
- [15] J. D. Bain, E. S. Diala, C. G. Glabe, T. A. Dix, A. R. Chamberlin, *J. Am. Chem. Soc.* **1989**, *111*, 8013–8014.
- [16] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751–755.
- [17] T. Kawakami, H. Murakami, H. Suga, *Chem. Biol.* **2008**, *15*, 32–42.
- [18] A. O. Subtelny, M. C. T. Hartman, J. W. Szostak, *J. Am. Chem. Soc.* **2008**, *130*, 6131–6136.
- [19] H. Murakami, A. Ohta, H. Ashigai, H. Suga, *Nat. Methods* **2006**, *3*, 357–359.
- [20] Y. Goto, T. Katoh, H. Suga, *Nat. Protoc.* **2011**, *6*, 779–790.

- [21] T. Passioura, H. Suga, *Chem. Commun.* **2017**, 53, 1931–1940.
- [22] C. Tsiamantas, J. M. Rogers, H. Suga, *Chem. Commun.* **2020**, 56, 4265–4272.
- [23] N. S. van der Velden, N. Kälin, M. J. Helf, J. Piel, M. F. Freeman, M. Künzler, *Nat. Chem. Biol.* **2017**, 13, 833–835.
- [24] S. Ramm, B. Krawczyk, A. Mühlenweg, A. Poch, E. Mösker, R. D. Süßmuth, *Angew. Chem. Int. Ed.* **2017**, 56, 9994–9997; *Angew. Chem.* **2017**, 129, 10127–10130.
- [25] E. Matabaro, H. Kaspar, P. Dahlin, D. L. V. Bader, C. E. Murar, F. Staubli, C. M. Field, J. W. Bode, M. Künzler, *Sci. Rep.* **2021**, 11, 3541.
- [26] H. Song, N. S. van der Velden, S. L. Shiran, P. Bleiziffer, C. Zach, R. Sieber, A. S. Imani, F. Krausbeck, M. Aebi, M. F. Freeman, et al., *Sci. Adv.* **2018**, 4, eaat2720.
- [27] C. Ongpipattanakul, S. K. Nair, *ACS Chem. Biol.* **2018**, 13, 2989–2999.
- [28] H. Song, J. H. Naismith, *Curr. Opin. Struct. Biol.* **2020**, 65, 79–88.
- [29] H. Song, J. R. Fahrig-Kamaraukaitė, E. Matabaro, H. Kaspar, S. L. Shiran, C. Zach, A. Pace, B.-A. Stefanov, J. H. Naismith, M. Künzler, *ACS Chem. Biol.* **2020**, 15, 1901–1912.
- [30] M. R. Quijano, C. Zach, F. S. Miller, A. R. Lee, A. S. Imani, M. Künzler, M. F. Freeman, *J. Am. Chem. Soc.* **2019**, 141, 9637–9644.
- [31] M. Montalbán-López, T. A. Scott, S. Ramesh, I. R. Rahman, A. J. van Heel, J. H. Viel, V. Bandarian, E. Dittmann, O. Genilloud, Y. Goto, et al., *Nat. Prod. Rep.* **2021**, 38, 130–239.
- [32] F. Oldach, R. Al Toma, A. Kuthning, T. Caetano, S. Mendo, N. Budisa, R. D. Süßmuth, *Angew. Chem. Int. Ed.* **2012**, 51, 415–418; *Angew. Chem.* **2012**, 124, 429–432.
- [33] N. Budisa, *Curr. Opin. Biotechnol.* **2013**, 24, 591–598.
- [34] S. Zobel, S. Boecker, D. Kulke, D. Heimbach, V. Meyer, R. D. Süßmuth, *ChemBioChem* **2016**, 17, 283–287.
- [35] N. Kakkar, J. G. Perez, W. R. Liu, M. C. Jewett, W. A. van der Donk, *ACS Chem. Biol.* **2018**, 13, 951–957.
- [36] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 6705–6710.
- [37] M. Vila-Perelló, Z. Liu, N. H. Shah, J. A. Willis, J. Idoyaga, T. W. Muir, *J. Am. Chem. Soc.* **2013**, 135, 286–292.
- [38] H. Iwai, S. Züger, J. Jin, P.-H. Tam, *FEBS Lett.* **2006**, 580, 1853–1858.
- [39] J. Zettler, V. Schütz, H. D. Mootz, *FEBS Lett.* **2009**, 583, 909–914.
- [40] N. H. Shah, G. P. Dann, M. Vila-Perelló, Z. Liu, T. W. Muir, *J. Am. Chem. Soc.* **2012**, 134, 11338–11341.
- [41] N. H. Shah, T. W. Muir, *Chem. Sci.* **2014**, 5, 446–461.
- [42] A. J. Stevens, Z. Z. Brown, N. H. Shah, G. Sekar, D. Cowburn, T. W. Muir, *J. Am. Chem. Soc.* **2016**, 138, 2162–2165.
- [43] A. J. Stevens, G. Sekar, N. H. Shah, A. Z. Mostafavi, D. Cowburn, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2017**, 114, 8538–8543.
- [44] H. Liang, W. S. Sandberg, T. C. Terwilliger, *Proc. Natl. Acad. Sci. USA* **1993**, 90, 7010–7014.
- [45] H. L. Zhao, X. Q. Yao, C. Xue, Y. Wang, X. H. Xiong, Z. M. Liu, *Protein Expression Purif.* **2008**, 61, 73–77.
- [46] C. Rullán-Lind, R. B. Pietri, M. Vázquez-Cintrón, A. Baerga-Ortiz, *Protein Sci.* **2018**, 27, 969–975.

Manuscript received: January 18, 2021

Revised manuscript received: March 23, 2021

Accepted manuscript online: April 15, 2021

Version of record online: May 17, 2021