Check for updates

A route to diastereomerically pure phenylglycine thioester peptides: crucial intermediates for investigating glycopeptide antibiotic biosynthesis†

Julien Tailhades, [ID] [ab] Melanie Schoppet,[abc] Anja Greule,[ab] Madeleine Peschke,[c] Clara Brieke[c] and Max J. Cryle [ID] *[abc]

Non-ribosomal peptides contain an array of amino acid building blocks that can present challenges for the synthesis of important intermediates. Here, we report the synthesis of glycopeptide antibiotic (GPA) thioester peptides that retains the crucial stereochemical purity of the terminal phenylglycine residue, which we show is essential for the enzymatic GPA cyclisation cascade.

Non-ribosomal peptide synthesis is central to the biosynthesis of many compounds of medical importance, including a large number of antibiotics.[1–4] Due to the modular structure of non-ribosomal peptide synthetases (NRPSs) – comprising repeating domains performing specific catalytic functions – and their non-dependence on the ribosome, NRPS assembly lines are able to synthesise peptides from a wide range of amino acids.[1,3] This feature, combined with the extensive incorporation of (D)-amino acids and large range of further structural modifications, contribute to the extensive diversity of natural NRPS-peptides.[1,2] Given the complexity of many of these compounds that are of medical interest – such as the glycopeptide antibiotics (Fig. 1), with vancomycin and teicoplanin as representative members – there is a need to characterise NRPS biosynthesis in order to explore the possibilities for compound development through redesigning the corresponding biosynthetic machinery.[1–4] In this regard, GPAs are perfect examples: total synthesis is highly challenging, meaning that all GPAs in clinical use stem from bacterial production *in vivo* and modified GPAs generated *via* total synthesis are not always accessible for the use in clinics.[4,5] The main reasons for this synthetic complexity are the high content of racemisation-prone
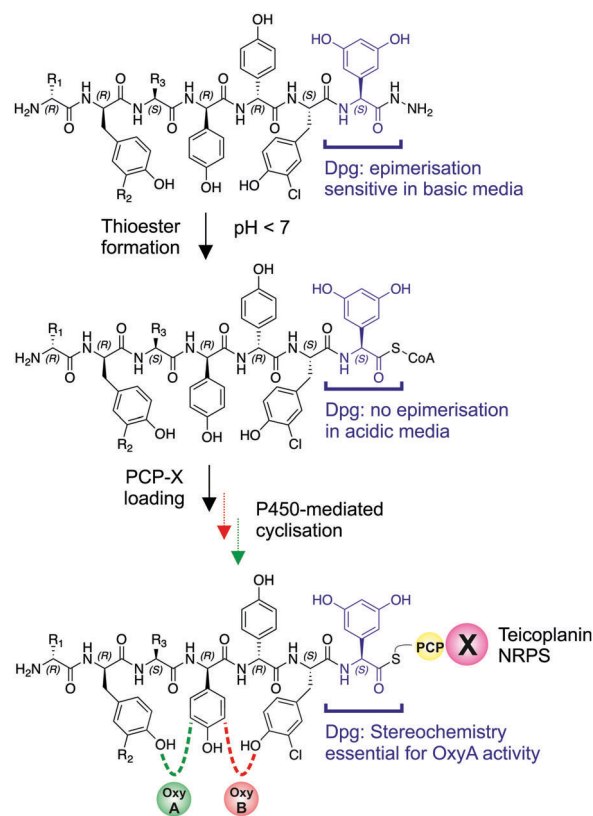

Fig. 1 GPA cyclisation *in vitro* is enabled by diastereomerically pure phenylglycine thioester peptides: teicoplanin seq. **5**: R_1 = 4-Hpg; R_2 = Cl; R_3 = 3,5-Dpg; pekiskomycin seq. **6**: R_1 = CH$_3$; R_2 = H; R_3 = Glu; actinoidin seq. **7**: R_1= 4-Hpg; R_2 = H; R_3 = Phe; vancomycin seq. **8**: R_1 = Leu; R_2 = Cl; R_3 = Asn. PCP – peptidyl carrier protein domain, X – Oxy recruitment domain, A_{tei} – OxyA from the teicoplanin system, B_{van} – OxyB from the vancomycin system.

[a] EMBL Australia, Monash University, Clayton, Victoria 3800, Australia.
 E-mail: max.cryle@monash.edu

[b] The Monash Biomedicine Discovery Institute, Department of Biochemistry and
 Molecular Biology and ARC Centre of Excellence in Advanced Molecular Imaging,
 Monash University, Clayton, Victoria 3800, Australia

[c] Department of Biomolecular Mechanisms, Max Planck Institute for Medical
 Research, Jahnstrasse 29, 69120 Heidelberg, Germany

† Electronic supplementary information (ESI) available: Complete synthetic details and compound characterisation; enzymatic assays and traces. See DOI: 10.1039/c7cc09409d

phenylglycine residues as well as the highly crosslinked structure of GPAs (vancomycin: AB, C-*O*-D, D-*O*-E rings; teicoplanin AB, C-*O*-D, D-*O*-E, F-*O*-G rings), which is formed by cytochrome P450 (Oxy) enzymes that interact with the unique NRPS domain found in GPA biosynthesis, the X-domain.[4,6,7]

The potential redesign of NRPS machineries to produce novel compounds has long been recognised as highly desirable.[1,2] To achieve redesign, a comprehensive understanding of the NRPS assembly line is required in order to generate modified NRPS enzymes that retain efficiency and productivity for modified peptides. In this regard, the characterisation of the rates and specificities of individual domains – such as adenylation (A)-domains, responsible for amino acid selection and activation, condensation (C)-domains, responsible for peptide bond formation, epimerisation (E)-domains, responsible for peptide epimerisation, and thioesterase (TE)-domains, ultimately responsible for peptide cleavage from the NRPS – is key to ensuring that modified assembly lines will be functional.[1,2] Additionally, GPA biosynthesis requires a functional cyclisation cascade of Oxy enzymes in order to produce bioactive compounds, and thus the selectivity of Oxy enzymes in this cascade must be clarified before peptide redesign is undertaken.[6] In all these cases, however, the assessment of enzymatic activity is complicated by the necessity of peptide intermediates to be bound to peptidyl carrier protein (PCP)-domains, which serve as an attachment point for all amino acid and peptide intermediates during NRPS biosynthesis.[1,2,8] The ability to enzymatically load coenzyme A (CoA) substrates onto PCP domains using the promiscuous phosphopantetheinyl transferase Sfp has made a vital contribution to overcome this problem, as it allows biosynthetic steps to be interrogated without complete reconstitution of the NRPS.[1,8] It does, however, require effective methods for the generation of peptide thioester CoA substrates to be able to undertake these experiments. However, the diversity of amino acid monomers utilised by NRPS machineries can now present serious problems for such syntheses, with no better example found than the phenylglycine residues that form the majority of residues in the peptide core of GPAs such as teicoplanin.[3] Phenylglycine residues, by virtue of their structure, are highly epimerisation prone residues, especially under the basic conditions required for Fmoc-based solid phase peptide synthesis (SPPS) strategies.[3] Through the careful selection of activation and coupling conditions, it is possible to generate peptides (such as GPA precursor peptides) using Fmoc synthesis.[9–11] However, the requirement to synthesise peptide CoA species for PCP loading adds a further complication, as the generation of the C-terminal thioester necessitates either activation and coupling or thioester exchange, both of which have high potential to racemise the C-terminal residue. This is particularly problematic in the case of GPAs (the terminal Dpg residue)[4] or β-lactam antibiotics such as nocardicin (the terminal Hpg residue), where the stereochemistry of this residue plays an important role in subsequent biosynthesis steps.[12] In GPA biosynthesis, we have recently demonstrated the necessity of maintaining the correct stereochemistry of C-terminal phenylglycine residues, where the activity of the second essential P450 in the GPA cyclisation cascade – OxyA – was shown to depend on the (L)-configuration of the terminal residue.[13–16] As complete racemisation of this residue occurred under previous Fmoc-synthetic conditions,[10,11] access to enantiomerically pure peptides was restricted and based purely on the ability to affect their chromatographic separation. Whilst resolution of simplified GPA peptides – those containing a C-terminal Hpg residue – was possible, resolution of those containing the natural Dpg residue

was not. The C-terminal Dpg residue in GPAs is part of the AB ring, which is crucial for GPA activity. This crosslink requires Dpg to generate the correct ring size (a Hpg-containing GPA has no antibiotic activity),[17,18] making the Dpg residue essential for any exploitation of the GPA cyclisation cascade for antibiotic development. Thus, we sought to develop a method to generate peptide thioesters that maintains the stereochemical purity of C-terminal residues – in spite of their propensity for racemisation – as probes of NRPS function, with a specific focus on GPA biosynthesis.

The instability of the thioester linkage in Fmoc-based SPPS has led to significant efforts to enable their synthesis.[19] Amongst these strategies,[20] we focused on approaches that were straightforward, efficient and achievable in acidic media.[21,22] Indeed, epimerisation is often triggered under basic conditions when the C-terminal amino acid is not urethane-protected, which is the case during thioester preparation. Peptide hydrazides as precursors of peptide thioesters appeared to be excellent strategies since these did not require any specific linkers.[21,22] Additionally, the C-terminal intermediate acyl azide can be directly converted into thioester without the conversion into a more stable thioester linkage such as MPAA or TFET, which is normally needed for native chemical ligation.[10,11,23]

Our first priority was to validate the conversion of the peptide hydrazides into peptide CoA thioesters to determine whether we could access peptide CoAs with a high enantiomeric purity. To this end, three tripeptide hydrazides that mimic the natural terminus of GPA precursors (X_7-Cl-Tyr_6-D-Hpg_5) incorporating either a C-terminal L-3,5-hydroxyphenylglycine (L-Dpg, 1a), L-4-hydroxyphenylglycine (L-Hpg, 2b) and L-phenylglycine residue (L-Phg, 3c) were prepared using our optimised protocol, with the use of DBU for Fmoc removal and COMU/2,6-lutidine for Fmoc-amino acid coupling (Scheme 1).‡[11,16] Conversion of the peptide hydrazides 1a–3a into their corresponding peptide CoA thioesters 1–3 was subsequently achieved in 1–2 hours and in quantitative yield (Scheme 1).

All purified peptide hydrazides 1a–3a and CoA thioesters 1–3 were characterised by ^{1}H NMR spectroscopy (SI2, ESI†) and no
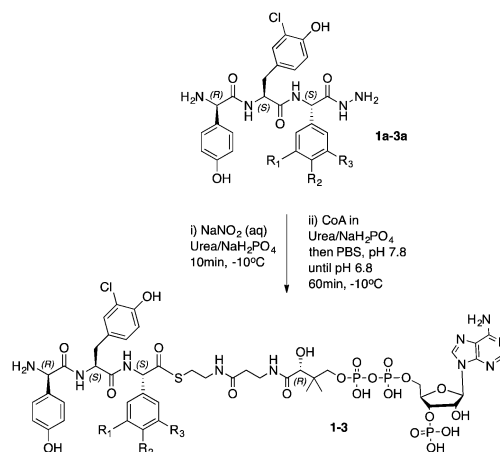
Scheme 1 Synthesis of C-terminal coenzyme A tripeptide thioesters through the peptide hydrazide strategy to avoid epimerisation: L-Dpg (1a/1): $R_1 = R_3 = -OH$, $R_2 = -H$; L-Hpg (2a/2): $R_1 = R_3 = -H$, $R_2 = -OH$; L-Phg (3a/3): $R_1 = R_2 = R_3 = -H$.

Table 1 Synthesis and enzymatic turnover of GPA heptapeptides **4–8**

#	Sequence	Yield (%)	Monocyclic (%)	Bicyclic (%)	Rel. OxyA$_{tei}$ (%)
4	Tei–CoA (*rac* 7-Dpg)	<2	20 ± 1	1 ± 0.1	4 ± 2
5	Tei–CoA	7	79 ± 2	42 ± 7	52 ± 8
6	Pek–CoA	27	86 ± 4	61 ± 11	71 ± 11
7	Act–CoA	13	28 ± 8	15 ± 5	53 ± 6
8	Van–CoA	15	40 ± 7	12 ± 2	29 ± 7

epimerisation of the critical thioester Hα was detected for the sequences **2** and **3**; epimerisation was below 5% for **1**. With this effective route in hand, we then turned to the synthesis of heptapeptide CoAs. Our initial target was the complete teicoplanin precursor peptide, containing two chlorinated tyrosine residues as well as both L-Dpg residues at positions 3 and 7 of the peptide. The heptapeptide hydrazide **5a** was synthesised and converted into the corresponding peptide CoA thioester **5** in order to investigate the acceptance of this peptide by the GPA Oxy cyclisation cascade. For comparison, we also tested the acceptance of a racemic version of a comparable peptide **4** (only different in the presence of a Hpg residue at position 3 of the peptide). We then tested the level of cyclisation seen in these peptides using the proven enzymatic coupling of OxyB$_{van}$ and OxyA$_{tei}$,[7,13–15,24,25] with the peptides themselves loaded onto a PCP-X di-domain construct excised from the final module of the teicoplanin NRPS machinery.[7]

The results of the Oxy-catalysed turnover of peptides **4–5** demonstrated how vital the L-configuration of the C-terminal residue is for effective enzymatic crosslinking. For the racemic peptide **4**, monocyclisation was now seriously impaired by the

presence of the incorrect peptide diastereomer, which reduced the levels of monocyclisation to below 20% (Table 1). In contrast to this, the oxidative cyclisation of the pure peptide diastereomer **5** by OxyB$_{van}$ remained highly effective (∼80%). However, the effect of peptide stereochemistry was more extreme when investigating the installation of the second crosslink by OxyA$_{tei}$ (SI4, ESI†): bicyclisation of **5** remained effective (∼50%), whilst the racemic peptide **4** was barely crosslinked by this enzyme (<4%, Table 1 and Fig. 2D). To further demonstrate the utility of the hydrazide synthesis route, we synthesised heptapeptide hydrazides of GPAs pekiskomycin (**6a**), actinoidin (**7a**), and vancomycin (**8a**). These were subsequently converted to their corresponding peptide CoA thioesters **6–8** and loaded onto the PCP-X di-domain to investigate the Oxy acceptance of these heptapeptides. Impressive results were obtained with the pekiskomycin sequence **6**, which led to the total formation of >60% of bicyclic peptide representing relative OxyA$_{tei}$ activity of 71% (Fig. 2B). The activity of this peptide in the linear, mono- and bicyclic forms was assessed using zone inhibition assays, which confirmed that these peptides did not have appreciable anti-microbial activity (SI5, ESI†). Cyclisation of the actinoidin sequence **7** was limited by precipitation during PCP loading, which has been previously observed for certain PCP-X/peptide pairings and is attributable to the hydrophobic nature of **7**.[13] Nonetheless OxyA activity remained above 50% for **7** (SI4, ESI†). The vancomycin sequence **8** was cyclised by both enzymes, albeit at lower levels, which indicates that further testing of PCP-X/peptide/Oxy combinations will be needed in future to maximise the peptide cyclisation efficiency of the cascade in specific cases (Fig. 2C and D).
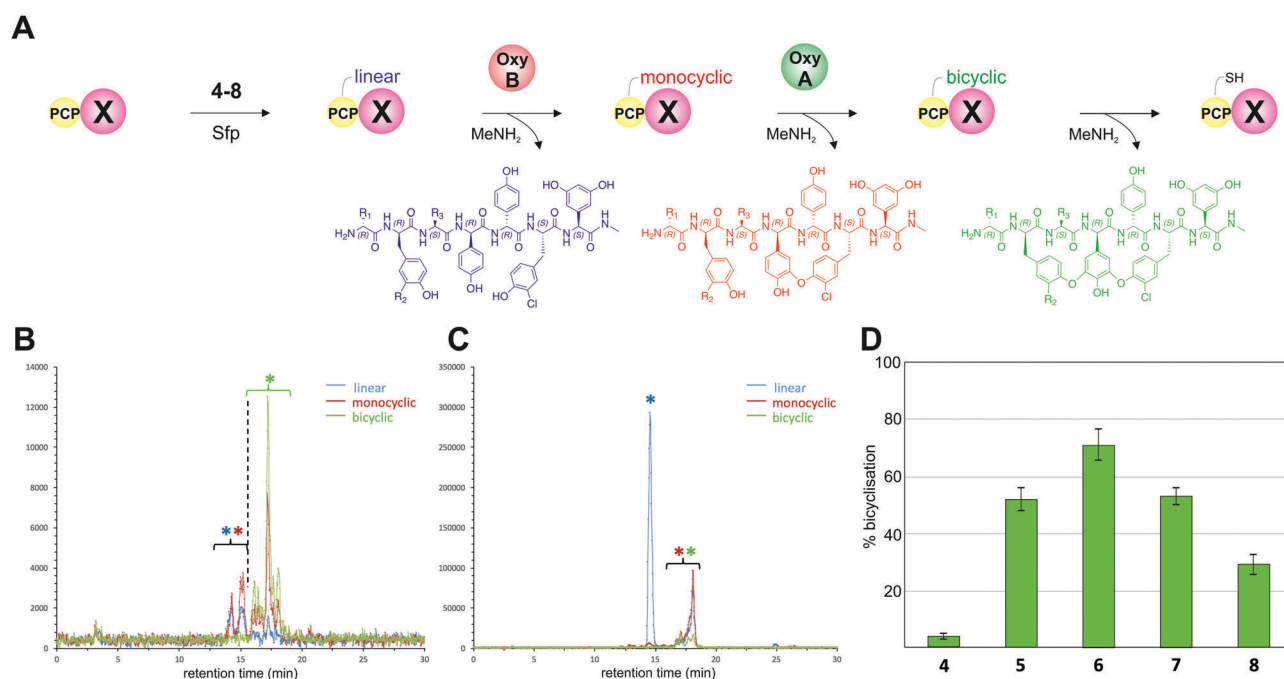

Fig. 2 Enzymatic crosslinking of heptapeptides **4–8**. (A) Schematic representation of the turnover protocol starting with the loading of peptide CoAs onto the PCP-X di-domain, followed by the sequential peptide crosslinking catalysed by OxyB$_{van}$ and OxyA$_{tei}$.; (B and C) HPLC-MS analysis of linear, monocyclic and bicyclic pekiskomycin (**6**) and vancomycin (**8**) sequences, X-axis – time (minutes), Y-axis – MS intensity; (D) summary of OxyA$_{tei}$ activity against peptides with standard deviation (3 replicates) **4–8**.

The ability to generate GPA peptides that maintain the stereochemical purity of the C-terminal Dpg residue will allow such studies to be undertaken now. Even more importantly, this route opens the door to future efforts to investigate the installation of the AB crosslink into these complex peptide natural products. Our synthetic strategy will also prove highly useful for the investigation of the many novel NRPS and NRPS/PKS systems that produce important phenylglycine containing peptide products.[3,26–28]

Conflicts of interest

There are no conflicts to declare.

Notes and references

‡ Briefly, peptide synthesis was carried out on a Protein Technologies Tribute synthesizer: Fmoc removal was carried out with a 1% (v/v) DBU solution in DMF using UV feedback monitoring; coupling was performed by activation of Fmoc-amino acids (3 eq.) in the presence of COMU (3 eq.) and 2,6-lutidine (3 eq.). Protecting groups and resin were cleaved using a solution of TFA/TIS/H_2O (95/2.5/2.5, v/v/v) for 1 h at room temperature. After cleavage, the resin was removed by filtration and washed twice with TFA, with the filtrate concentrated under a stream of nitrogen until $\sim$2 mL volume. The peptide products were precipitated in ice cold Et_2O and washed by centrifugation three times. After purification by preparative RP-HPLC, the peptide hydrazide was dissolved in 6 M urea buffer containing 0.2 M of NaH_2PO_4 (pH 3, see ESI†) to give a final peptide concentration of 4–5 mM. The temperature of the reaction was maintained between -10 °C and -15 °C using a mixture of ice and sodium chloride. The addition of a 0.5 M $NaNO_2$ in water (0.95 eq.) allowed the formation of the acyl azide in $\sim$10 minutes. Then, CoA (1.2 eq.) dissolved in 6 M urea buffer containing 0.2 M of NaH_2PO_4 (pH 3, see ESI†) was added dropwise followed by a solution of 1 M potassium phosphate buffer (pH 8, see ESI†) until a pH of 6.5–6.8 was obtained. Reaction monitoring was performed every 30 minutes using LC/MS, with purification performed by preparative RP-HPLC after completion of the reaction (1–2 hours).

Whilst our study was ongoing a report of GPA peptide synthesis using a hydrazide method was reported;[29] however, this peptide contained a non-standard, racemisation-resistant tyrosine residue at the peptide C-terminus.

1 R. D. Süssmuth and A. Mainz, *Angew. Chem., Int. Ed.*, 2017, **56**, 3770–3821.
2 J. A. E. Payne, M. Schoppet, M. H. Hansen and M. J. Cryle, *Mol. BioSyst.*, 2017, **13**, 9–22.
3 R. S. Al Toma, C. Brieke, M. J. Cryle and R. D. Suessmuth, *Nat. Prod. Rep.*, 2015, **32**, 1207–1235.
4 G. Yim, M. N. Thaker, K. Koteva and G. Wright, *J. Antibiot.*, 2014, **67**, 31–41.
5 A. Okano, N. A. Isley and D. L. Boger, *Chem. Rev.*, 2017, **117**, 11952–11993.
6 M. Peschke, M. Gonsior, R. D. Süssmuth and M. J. Cryle, *Curr. Opin. Struct. Biol.*, 2016, **41**, 46–53.
7 K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch and M. J. Cryle, *Nature*, 2015, **521**, 105–109.
8 T. Kittilä, A. Mollo, L. K. Charkoudian and M. J. Cryle, *Angew. Chem., Int. Ed.*, 2016, **55**, 9834–9840.
9 C. Liang, M. A. M. Behnam, T. R. Sundermann and C. D. Klein, *Tetrahedron Lett.*, 2017, **58**, 2325–2329.
10 C. Brieke, V. Kratzig, K. Haslinger, A. Winkler and M. J. Cryle, *Org. Biomol. Chem.*, 2015, **13**, 2012–2021.
11 C. Brieke and M. J. Cryle, *Org. Lett.*, 2014, **16**, 2454–2457.
12 N. M. Gaudelli and C. A. Townsend, *Nat. Chem. Biol.*, 2014, **10**, 251–258.
13 M. Peschke, C. Brieke, R. J. A. Goode, R. B. Schittenhelm and M. J. Cryle, *Biochemistry*, 2017, **56**, 1239–1247.
14 V. Ulrich, C. Brieke and M. Cryle, *Beilstein J. Org. Chem.*, 2016, **12**, 2849–2864.
15 M. Peschke, K. Haslinger, C. Brieke, J. Reinstein and M. Cryle, *J. Am. Chem. Soc.*, 2016, **138**, 6746–6753.
16 C. Brieke, M. Peschke, K. Haslinger and M. J. Cryle, *Angew. Chem., Int. Ed.*, 2015, **54**, 15715–15719.
17 D. Bischoff, B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G. J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben and R. D. Suessmuth, *ChemBioChem*, 2005, **6**, 267–272.
18 S. Weist, C. Kittel, D. Bischoff, B. Bister, V. Pfeifer, G. J. Nicholson, W. Wohlleben and R. D. Süssmuth, *J. Am. Chem. Soc.*, 2004, **126**, 5942–5943.
19 A. B. Clippingdale, C. J. Barrow and J. D. Wade, *J. Pept. Sci.*, 2000, **6**, 225–234.
20 J. Tailhades, N. A. Patil, M. A. Hossain and J. D. Wade, *J. Pept. Sci.*, 2015, **21**, 139–147.
21 G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui and L. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 7645–7649.
22 J.-S. Zheng, S. Tang, Y.-K. Qi, Z.-P. Wang and L. Liu, *Nat. Protoc.*, 2013, **8**, 2483.
23 K. Woithe, N. Geib, K. Zerbe, D. B. Li, M. Heck, S. Fournier-Rousset, O. Meyer, F. Vitali, N. Matoba, K. Abou-Hadeed and J. A. Robinson, *J. Am. Chem. Soc.*, 2007, **129**, 6887–6895.
24 V. Ulrich, M. Peschke, C. Brieke and M. Cryle, *Mol. BioSyst.*, 2016, **12**, 2992–3004.
25 M. Peschke, C. Brieke and M. J. Cryle, *Sci. Rep.*, 2016, **6**, 35584.
26 L. Ray, K. Yamanaka and B. S. Moore, *Angew. Chem., Int. Ed.*, 2016, **55**, 364–367.
27 M. Gonsior, A. Mühlenweg, M. Tietzmann, S. Rausch, A. Poch and R. D. Süssmuth, *ChemBioChem*, 2015, **16**, 2610–2614.
28 W.-T. Liu, R. D. Kersten, Y.-L. Yang, B. S. Moore and P. C. Dorrestein, *J. Am. Chem. Soc.*, 2011, **133**, 18010–18013.
29 C. C. Forneris, S. Ozturk, M. I. Gibson, E. J. Sorensen and M. R. Seyedsayamdost, *ACS Chem. Biol.*, 2017, **12**, 2248–2253.