COMMUNICATION

A chemoenzymatic approach to the synthesis of glycopeptide antibiotic analogues

Julien Tailhades,*[a,b] Yongwei Zhao, [a,b] Candace Ho, [a,b] Anja Greule, [a,b] Iftekhar Ahmed,[c] Melanie Schoppet,[a,b] Ketav Kulkarni,[a] Rob J. A. Goode,[a,d] Ralf B. Schittenhelm,[a,d] James J. De Voss[c] and Max J. Cryle*[a,b]

[a] Dr. J. Tailhades, Y. Zhao, C. Ho, Dr. A. Greule, Dr. M. Schoppet, Dr. K. Kulkarni, R.J.A. Goode, Dr. R.B. Schittenhelm, A/Prof. Dr. M.J. Cryle
Department of Biochemistry and Molecular Biology
Monash Biomedicine Discovery Institute
Monash University, Clayton, Victoria 3800, Australia
E-mail: julien.tailhades@monash.edu and max.cryle@monash.edu

[b] Dr. J. Tailhades, Y. Zhao, C. Ho, Dr. A. Greule, Dr. M. Schoppet, A/Prof. Dr. M.J. Cryle
EMBL Australia
Monash University, Clayton, Victoria 3800, Australia

[c] Dr. I. Ahmed, Prof. Dr. J.J. De Voss
School of Chemistry and Molecular Biosciences
The University of Queensland
St Lucia, Queensland 4072, Australia

[d] R.J.A. Goode, Dr. R.B. Schittenhelm
Monash Proteomics & Metabolomics Facility
Department of Biochemistry and Molecular Biology
Monash Biomedicine Discovery Institute
Monash University, Clayton, Victoria 3800, Australia

Supporting information for this article is given via a link at the end of the document.

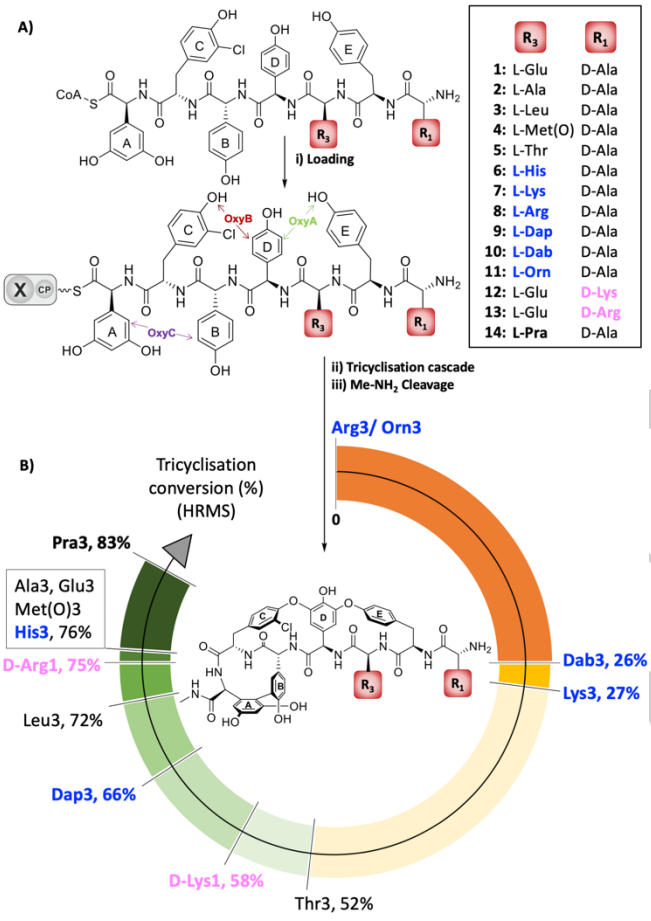
Abstract: The glycopeptide antibiotics (GPAs) are important antibiotics that are highly challenging to synthesise due to their unique and heavily crosslinked structure. Given this, the synthetic production and diversification of this key compound class remains impractical. Furthermore, the possibility of biosynthetic reengineering of GPAs is not yet feasible as the selectivity of the biosynthetic crosslinking enzymes for altered substrates is largely unknown. Here, we show that the combination of peptide synthesis with enzymatic cyclisation enables the formation of novel examples of GPAs and provides an indication of the utility of these crucial enzymes. By accessing the biosynthetic process *in vitro*, we identify peptide modifications that are enzymatically tolerated and can also reveal the mechanistic basis for substrate intolerance where present. Using this approach, we next specifically activate modified residues within GPAs for functionalisation at previously inaccessible positions, offering the possibility of late-stage chemical functionalisation after GPA cyclisation is complete.

Complex natural products, elegantly biosynthesised, have provided a wealth of bioactive compounds that we as a society exploit, particularly in medicine. Here, the products resulting from bacterial competition over millions of years' evolution have led to a wealth of intricate structures with a variety of antibiotic activities.[1] Whilst synthetic chemistry has played a vital role in the structural elucidation and diversification of many important compound classes, the complexity of the products of large assembly lines – the polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) – is often sufficiently high to prevent effective total synthesis at a scale viable for medical use.[2] In this regard, the glycopeptide antibiotics (GPAs) – exemplified by the clinical compounds vancomycin and teicoplanin – represent a potent example of a class of compounds with such challenging structures that all compounds used from this class, be they natural or semi-synthetically are derived from naturally produced GPAs.[3] The recent identification of divergent GPAs with a new mode of action – the inhibition of bacterial cell wall remodelling by autolysins – offers yet further possibilities for development of this complex compound class.[4]

In the biosynthesis of all GPAs, the non-ribosomal synthesis of heptapeptide precursors is immediately followed by a complex cyclisation cascade mediated by a number of cytochrome P450 (Oxy) enzymes.[2, 5] Within the cascade found in lipid II binding GPAs, each Oxy enzyme is responsible for the insertion of one crosslink into the final GPA, with a defined order of activity commencing with OxyB installing the C-*O*-D crosslink, followed by OxyA installing the D-*O*-E crosslink and finally OxyC installing the AB ring. This multi-step cascade – in which recruitment of the P450s to the NRPS-bound peptide is mediated by a unique recruitment domain (the X-domain) – is responsible for the generation of the three dimensional shape of GPAs.[5] This is in turn required for their ability to bind and sequester lipid II, the source of antibiotic activity for typical GPAs.[3a, 3b] Comparable crosslinks are also found in divergent GPAs, although in these cases certain Oxy enzymes can install multiple rings in the final crosslinked peptide.[4, 6] Given that these cyclisation steps are the origin of both the activity of GPAs and their synthetic challenge,[7] significant research has focussed on the possibility of exploiting a chemoenzymatic cascade for the synthesis and possible diversification of GPA structures.[8]

Previous results have shown the potential to generate significant conversion of synthetic linear peptides into tricyclic species using the Oxy enzymes ($OxyB_{van}$, $OxyA_{ris}$ and $OxyC_{cep}$), albeit with

limited exploration of the biocatalytic capabilities of these enzymes.[8a, 9] In this work, we now perform a comprehensive analysis of the tolerance of this enzymatic cascade for alternations in the structure of the precursor peptide in order to examine the general versatility of this chemoenzymatic approach. In doing so, we exploited the abilities of this methodology to explore the targeted installation of functionality within such GPA peptides, in particular with residues capable of selective, late stage functionalisation through the use of click chemistry. The results obtained support the general utility of these enzymes as functional biocatalysts, and highlight the importance of combining chemical and biosynthetic modifications in the future redesign of GPAs.



Scheme 1. GPA peptide tricyclisation cascade workflow (A) and representation of tricyclic conversion (B) from lowest (orange) to highest (green) as calculated by HRMS. (A(i)) 40 μM Peptide-CoA (**1-14-CoA**), 80 μM PCP-X_{tei}, 4 μM Sfp R4-4 mutant, 50 mM HEPES pH 7.0, 50 mM NaCl, 10 mM $MgCl_2$, 1 h, 30°C. (A(ii)) 0.5 μM OxyB, 0.66 μM PuR, 2.5 μM PuxAV+ mutant, 2 mM NADH, 33 μg/ mL glucose dehydrogenase, 0.3% w/v glucose, 2.5 mM TCEP, 0.25 M cysteamine, 15 min, 30°C; then 1 μM OxyA, 15 min, 30°C; 1 μM OxyC, 15 h, 30°C. (A(iii)) Methylamine, 15 min, RT.

Cyclisation of diverse peptide precursors is supported by GPA enzymes. With the ability to generate tricyclic peptides with high conversion via our chemoenzymatic synthesis route for both vancomycin and teicoplanin,[9] we first examined the ability of such a cascade to generate GPAs bearing different residues at position 3 of the peptide, with the rational being that the synthesis of modified GPAs at this position is highly challenging. Having previously observed successful cyclisation of precursor peptides bearing either carboxylic acid (pekiskomycin), amide (vancomycin) or aromatic (teicoplanin) moieties at position 3, we now explored the acceptance of 11 further GPA peptides containing altered residues at this position (**Scheme 1A**). In these experiments, we observed that conversion of precursor peptides containing either aliphatic residues alanine (**2**)/ leucine (**3**) or polar residues (methionine-oxide (**4**)/ threonine (**5**) was highly effective, indicating a high degree of substrate tolerance for the chemoenzymatic cascade for these modifications. As has been observed previously, incorporation of small molecules capable of protecting the Oxy enzymes from oxidative damage (TCEP and cysteamine) were required in all cases to achieve high level conversion, with the conversion level observed without such small molecules inversely correlated with the hydrophilicity of the residue incorporated at position 3 of the peptide. The only class of substrates where effective tricyclisation was not always maintained was for positively charged residues, for whilst a histidine-containing peptide (**6**) was well accepted, lysine (**7**) was less well accepted and arginine (**8**) not accepted at all, although in the case of **8** significant peptide hydrolysis from the carrier protein was observed and likely contributed to the lack of tricyclisation observed here (**Scheme 1B** and **SI Figures 1-8**).

Positive charged residues within GPAs can alter cyclisation efficiency. Given the decrease in activity for peptides **7-8**, we next examined the general acceptance of positively charged residues within GPA peptides using the chemoenzymatic cascade. First, we synthesised three further peptides containing chain-shortened lysine equivalents at position 3 of the peptide chain; specifically, Dap (**9**), Dab (**10**) and Orn (**11**). Tricyclisation conversion of these peptides showed significant variation depending on the chain length of the residue, with **10-11** poorly accepted by the cascade (**Scheme 1B** and **SI Figures 9-13**). However, the acceptance of **9** at twice the levels of conversion for **10** showed that the effective incorporation of alkylamine side chains is possible, and highlights the value of being able to screen for Oxy enzyme acceptance of different peptide substituents prior to engaging in biosynthetic redesign of the NRPS machinery that produces such peptides *in vivo*.

We determined that the reduction in cyclisation for **7**, **10** and **11** was the result of decreasing activity of both OxyB and OxyA enzymes but not OxyC (**Figure 1A**, above 70% of conversion from bicyclic to tricyclic mediated by OxyC), which suggests that the presence of the alkylamine residue immediately adjacent to residue 4 (which is common to the crosslinks installed by both OxyA and OxyB) is the cause of the low yields in these cases. Further support for this hypothesis was gained from the high activity of the chemoenzymatic cascade towards peptides containing D-lysine (**12**) or D-arginine (**13**) in position 1 of the peptide, indicating that it is not merely the presence of a positively charged residue that causes the inhibition of Oxy-catalysed cyclisation activity, but rather the ability of this specific residue to interfere with the activity of the Oxy enzymes themselves (**Figure 1A** and **SI Figures 12-13**).

To further understand the loss of OxyB or OxyA activity for **11**, the binding of peptidyl-PCP-X substrates to the Oxy enzymes was measured by UV spectroscopy (**Figure 1B-C** and **SI Figures 15-16**). By measuring the change in spin state of the heme iron with increasing concentration various peptidyl-PCP-X substrates versus OxyB or OxyA, we confirmed that the change in spin state

for poorly processed substrates (OxyB/ linear **11**-PCP-X (Orn3 – linear) and OxyA/ monocyclic **7**-PCP-X (Lys3 – monocyclic)) was greatly reduced compared to pairings known to be highly productive (OxyB/ linear **7**-PCP-X (Lys3 – linear) and OxyA/ monocyclic **14**-PCP-X (Pra3 – monocyclic)). Whilst all substrates show low (1-20 μM) micromolar interaction affinities for the Oxy enzymes, in agreement with previous measurement,[5, 8b] closer inspection of the difference spectra obtained from these measurements shows that effective P450 activation is only observed for substrates that show high levels of conversion (**SI Figures 15-16.** [10]

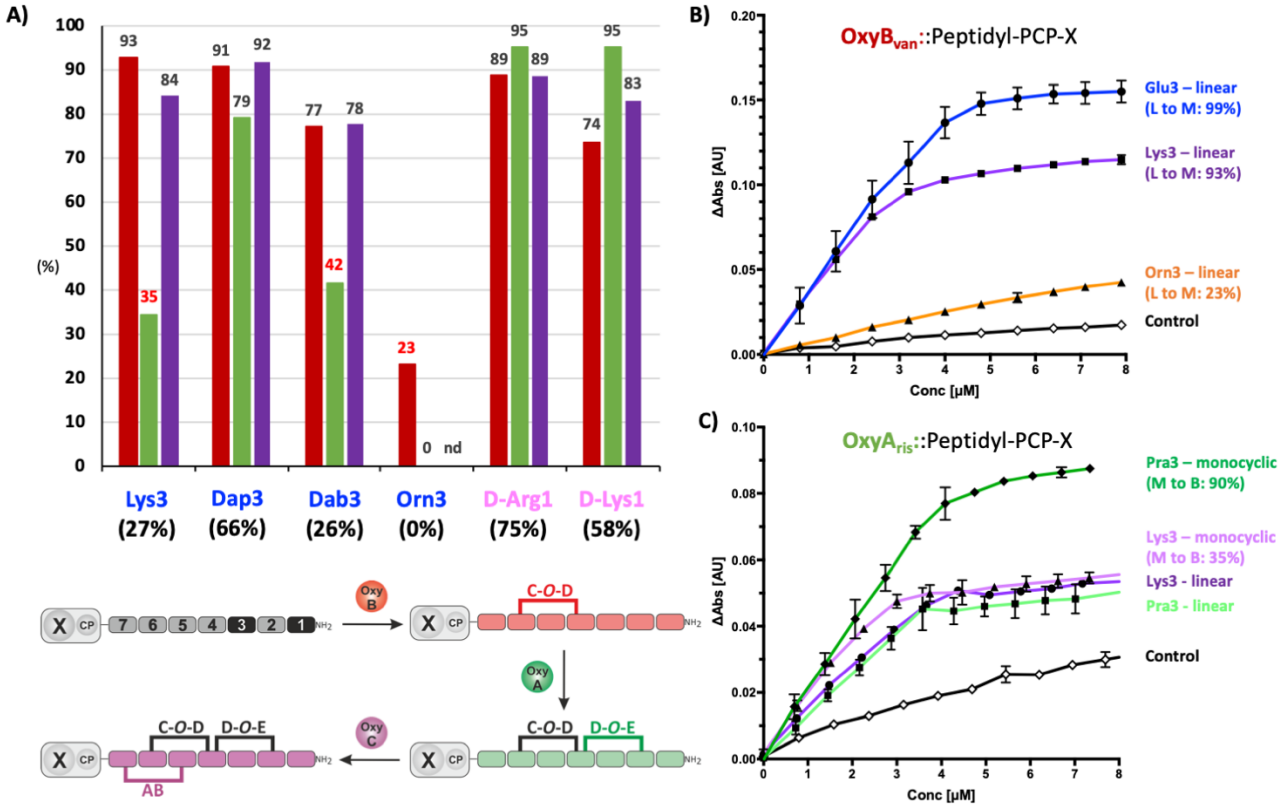


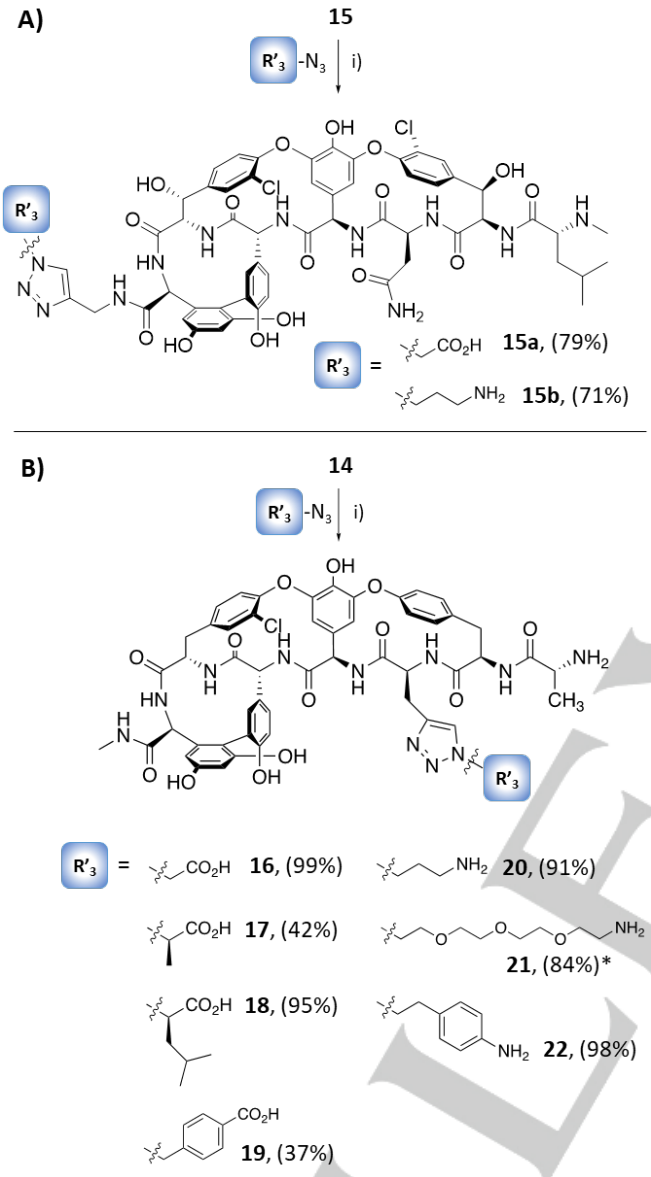
Figure 1. Cyclisation of peptides **7-13** showing conversion for individual Oxy enzymes as determined by HRMS (A) with a schematic description shown below and binding response of OxyB (B) and OxyA (C) measured by UV/Vis spectroscopy to PCP-X loaded peptides **1**, **7**, **11** and **14**. PCP loading was performed as indicated in Scheme 1 A(i). Oxy activities indicate conversion at each step (100% being total conversion of available peptide precursor for the specific enzyme into the product of that enzyme) and allows for identification of the loss of specific Oxy activity: OxyB-catalysed insertion of the C-*O*-D ring in the linear peptide shown in red, subsequent OxyA-catalysed insertion of the D-*O*-E ring shown in green and final OxyC-catalysed insertion of AB ring shown in purple (see panel (A) lower section). Overall conversion of linear peptide starting material into tricyclic peptide product is shown under the peptide name in brackets. Substrate binding curves shown in B) and C) allow determination of substrate affinity and the activation of Oxys by different substrates. Controls indicated are *apo*-PCP-X protein with no peptide loaded. Conversions for each of the peptides shown in the binding traces are indicated to the right of the graph (L – linear peptide; M – monocyclic peptide; B-bicyclic peptide).

In the case of OxyA binding, linear substrates **7**, **14** and **monocyclic-7**, which are not accepted as substrates, show red shifted absorption maxima that is indicative of substrates coordinating to, rather than displacing the water bound to the heme iron in the P450 resting state (**SI Figure 16**).[10] The ability to dissect the underlying cause of low peptide cyclisation activity by specific Oxy enzymes provides important clues into general substrate acceptance by the GPA cyclisation cascade. Given the generally poor acceptance of alkylamine residue at position 3 of the GPA peptide by OxyB/A, we next turned to late-stage diversification chemistry as a potential means to overcome these natural limitations of the GPA biosynthetic machinery.

Coupling enzymatic tricyclisation with late-stage GPA chemical diversification. Methodology for late-stage peptide modifications of GPAs has previously been developed for natural aglycones, opening the door to diversification of these complex structures after biosynthesis is complete.[11] The effective cyclisation of peptide **14** containing propargylglycine (Pra) at residue 3 provided us with the perfect scaffold with which to test the diversification of modified GPA peptides through such late-stage modification, as the Pra-alkyne would allow subsequent diversification with various azides via copper-catalysed click chemistry.[12] Initially, the modification of **15** (propargylamide vancomycin aglycone obtained through chemical synthesis) was achieved by click-chemistry as an optimisation template to determine reaction conditions for click chemistry on GPA aglycones. Modification of **15**-tricyclic using azidoacetic acid (**15a**) and azidopropanamine (**15b**) was explored, which afforded effective modification with conversions above 70% in both cases after RP-HPLC purification (**Scheme 2A** and **SI Figure 19-21**).

We subsequently scaled the tricyclisation reaction to allow the isolation of sufficient quantities of **14** to allow NMR analysis (**SI**

Figure 17-18). This analysis confirmed that the isolated peptide conformed to the anticipated aglycone structure, although the complexity of the spectrum suggested that a number of diastereomers and atropisomers were present, which prevented complete characterisation.



Scheme 2. Modification of tricyclic GPAs **14** and **15** via click chemistry. (A) Preliminary optimisation were performed using propargylamide vancomycin aglycone **15**. (i) R'3-azide, $CuSO_4$, sodium ascorbate, $NaHCO_3$, ACN/ H_2O (30:70), 2 h, RT. (B) Application of optimised derivatisation condition to the functionalisation of **14** (conversion determined by HRMS). *Major ion displays -2Da difference with theoretical MW (see **SI Figure 25**).

To demonstrate the utility of the Pra containing peptide **14** and coupling post-tricyclisation modification with the chemoenzymatic cascade, we modified **14** with 7 different azides to afford compounds **16-22** (**Scheme 2B** and **SI Figure 22-28**). Importantly, the alkyne introduced at residue 3 of a GPA is reactive under the conditions tested, and thus we could generate the compounds **16-22** with conversion ranging between 37% and 99% as determined by HRMS as previously performed with enzymatic tricyclisation. Crucially, the introduction of an amino function at residue 3 of the GPA, which was problematic using the direct chemoenzymatic strategy (**7**, **8**, **10** and **11**), is facile when introduced using post-cyclisation synthesis. Indeed, this led to conversions of above 75% for short aliphatic (**20**), PEGylated (**21**) and aniline moieties (**22**). Thus, the value of peptides such as **14** lies in the ability to explore a wider chemical space than is able to be tolerated by the existent Oxy enzymes. Unlike **15**, this does not rely on reactive functionality to be present in the natural aglycone, showing that it is now possible to specifically activate functionalities on previously inaccessible residues. This is also combined with the ease of diversification of single GPA aglycones to enable the facile screening of diverse GPA libraries through chemical diversification. In this regard the incorporation of Pra is particularly significant, as a biosynthesis pathway for this amino acid has recently been identified that would be able to provide this substrate for a suitably reengineered NRPS assembly line *in vivo*;[13] work on this is currently underway in our laboratory.

Taken together, in this work we have taken major steps towards the development and exploration of the scope of new methodology for the chemoenzymatic biosynthesis of new GPA aglycones in a manner that is then consistent with large scale production *in vivo* through biosynthetic reengineering. The Oxy cascade from GPA biosynthesis displays a wide tolerance for altered amino acids at position 3 of the GPA peptide, whilst the ability to analyse the *in vitro* yields of peptides in a facile manner allowed us to overcome low yields for positively charged alkylamine residues by altering the alkyl chain length. Perhaps more significantly, we have been able to incorporate the alkyne-containing Pra residue into tricyclic GPA aglycones and to confirm their structure. Chemical diversification of such a tricyclic peptide using click chemistry demonstrates the advantage of using such an approach to not only overcome Oxy enzyme selectivity but also to allow the rapid diversification of such peptides using specific, late stage chemical modification at positions inaccessible in the natural GPA. Given the recent identification of the novel mode of action of structurally divergent GPAs, this makes approaches to explore biosynthetic reengineering all the more pressing.[4a] The results obtained here demonstrate the potential for GPA redesign and provide concrete lead compounds to explore via the redesign of GPA biosynthesis, a prequel to identifying new, active members of this important, multifaceted antibiotic class.

Acknowledgements

G. Stier (BZH-Heidelberg) for fusion protein vectors; S. Bell (University of Adelaide) for redox proteins; J. Yin (University of Chicago) for the R4-4 Sfp expression plasmid. This work was supported by Monash University, EMBL Australia, the National Health and Medical Research Council (APP1140619 to M.J.C.) and under the Australian Research Council's Discovery Projects funding scheme (project number DP170102220 to J.J.D.V. and M.J.C.)

Keywords: glycopeptide antibiotics • cytochrome P450 • biosynthesis • biocatalysis • click chemistry

[1] a) C. T. Walsh, T. A. Wencewicz, *J. Antibiot.* **2014**, *67*, 7-22; b) M. N. Thaker, W. Wang, P. Spanogiannopoulos, N. Waglechner, A. M. King, R. Medina, G. D. Wright, *Nat. Biotech.* **2013**, *31*, 922.

[2] R. D. Süssmuth, A. Mainz, *Angew. Chem., Int. Ed.* **2017**, *56*, 3770-3821.

[3] a) G. Yim, M. N. Thaker, K. Koteva, G. Wright, *J. Antibiot.* **2014**, *67*, 31-41; b) M. A. T. Blaskovich, K. A. Hansford, M. S. Butler, Z. Jia, A. E. Mark, M. A. Cooper, *ACS Infect. Diseases* **2018**, *4*, 715-735; c) E. Marschall, M. J. Cryle, J. Tailhades, *J. Biol. Chem.* **2019**, *294*, 18769-18783.

[4] a) E. J. Culp, N. Waglechner, W. Wang, A. A. Fiebig-Comyn, Y.-P. Hsu, K. Koteva, D. Sychantha, B. K. Coombes, M. S. Van Nieuwenhze, Y. V. Brun, G. D. Wright, *Nature* **2020**, *578*, 582-587; b) H.-T. Chiu, B. K. Hubbard, A. N. Shah, J. Eide, R. A. Fredenburg, C. T. Walsh, C. Khosla, *Proc. Nat. Acad. Sci. USA* **2001**, *98*, 8548-8553.

[5] K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch, M. J. Cryle, *Nature* **2015**, *521*, 105-109.

[6] A. Greule, T. Izoré, D. Iftime, J. Tailhades, M. Schoppet, Y. Zhao, M. Peschke, I. Ahmed, A. Kulik, M. Adamek, R. J. A. Goode, R. B. Schittenhelm, J. A. Kaczmarski, C. J. Jackson, N. Ziemert, E. H. Krenske, J. J. De Voss, E. Stegmann, M. J. Cryle, *Nat. Commun.* **2019**, *10*, 2613.

[7] a) A. Okano, N. A. Isley, D. L. Boger, *Chem. Rev.* **2017**, *117*, 11952-11993; b) C. M. Crane, D. L. Boger, *J. Med. Chem.* **2009**, *52*, 1471-1476; c) J. J. McAtee, S. L. Castle, Q. Jin, D. L. Boger, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1319-1322.

[8] a) C. C. Forneris, M. R. Seyedsayamdost, *Angew. Chem., Int. Ed.* **2018**, *57*, 8048-8052; b) M. Peschke, K. Haslinger, C. Brieke, J. Reinstein, M. Cryle, *J. Am. Chem. Soc.* **2016**, *138*, 6746-6753; c) C. Brieke, M. Peschke, K. Haslinger, M. J. Cryle, *Angew. Chem., Int. Ed.* **2015**, *54*, 15715-15719.

[9] a) J. Tailhades, Y. Zhao, M. Schoppet, A. Greule, R. J. A. Goode, R. B. Schittenhelm, J. J. D. Voss, M. J. Cryle, *Org. Lett.* **2019**, *21*, 8635−8640; b) Y. Zhao, R. J. A. Goode, R. B. Schittenhelm, J. Tailhades, M. J. Cryle, *J. Org. Chem.* **2019**, *85*, 1537-1547.

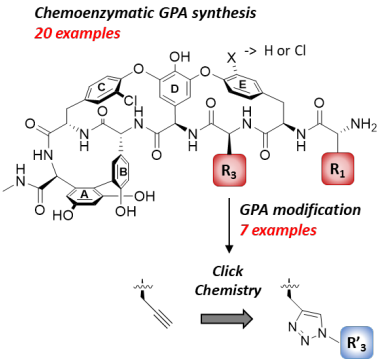
[10] A. Greule, J. E. Stok, J. J. De Voss, M. J. Cryle, *Nat. Prod. Rep.* **2018**, *35*, 757-791.

[11] a) T. P. Pathak, S. J. Miller, *J. Am. Chem. Soc.* **2013**, *135*, 8415-8422; b) T. P. Pathak, S. J. Miller, *J. Am. Chem. Soc.* **2012**, *134*, 6120-6123; c) D. T. Cohen, C. Zhang, C. M. Fadzen, A. J. Mijalis, L. Hie, K. D. Johnson, Z. Shriver, O. Plante, S. J. Miller, S. L. Buchwald, B. L. Pentelute, *Nat. Chem.* **2019**, *11*, 78-85.

[12] J. Zhang, J. Kemmink, D. T. S. Rijkers, R. M. J. Liskamp, *Org. Lett.* **2011**, *13*, 3438-3441.

[13] J. A. Marchand, M. E. Neugebauer, M. C. Ing, C. I. Lin, J. G. Pelton, M. C. Y. Chang, *Nature* **2019**, *567*, 420-424.

Entry for the Table of Contents



Glycopeptide antibiotic biosynthesis remains the only effective means to produce these important antibiotics. By investigating the final crosslinking steps of the biosynthesis *in vitro*, we can now show that the enzymes responsible for these challenging transformations possess broad substrate tolerance; poor enzymatic substrates can be overcome by using specific late-stage chemical modification of crosslinked peptides after enzymatic cyclisation.