



Cite this: DOI: 10.1039/c8ob02054j

Received 21st August 2018,
Accepted 8th October 2018

DOI: 10.1039/c8ob02054j

rsc.li/obc

Identification of a lysine 4-hydroxylase from the glidobactin biosynthesis and evaluation of its biocatalytic potential†

Alexander Amatuni[†] and Hans Renata^{†*}

We present the functional characterization of GlbB, a lysine 4-hydroxylase from the glidobactin biosynthetic gene cluster. Despite its narrow substrate specificity, GlbB is able to catalyze the hydroxylation of L-lysine with excellent total turnover number and complete regio- and diastereoselectivity. The synthetic utility of GlbB is illustrated by its use in the efficient preparation of a key dipeptide fragment of glidobactin.

The syrbactins are a family of hybrid peptide-polyketide natural products that share a 12-membered unsaturated macrolactam core (Fig. 1A).¹ Originally isolated as virulence factors in plant-pathogen interactions, the syrbactins were found to boost virulence by targeting the eukaryotic 20S proteasome core particle.² With reported IC₅₀ values of <20 nM, two members of this family, glidobactin A (1) and cepafungin I (2) represent some of the most potent natural proteasome inhibitors isolated to date.³ Crystallographic evidence suggests that inhibition occurs primarily by covalent binding of the α,β-unsaturated lactam to the chymotrypsin-like threonine residue in the β5-subunit,^{2–4} akin to the mode of action of several other proteasome inhibitors such as epoxomicin and salinosporamide A.⁵ Proteasome inhibitors have recently attracted attention as potential antitumor agents due to their ability to disrupt the regulated degradation of pro-apoptotic factors in malignant cells.⁶ At present, three proteasome inhibitors, namely bortezomib, carfilzomib, and ixazomib, have been approved by the FDA for the treatment of multiple myeloma, and a few others are undergoing clinical trials. Despite these developments, new proteasome inhibitors still need to be advanced due to the potential emergence of drug resistance and the syrbactins hold significant promise as such candidates. Towards this end, total syntheses efforts and subsequent medicinal chemistry development by Ichikawa,

Pirrung, and others have provided initial validation towards the use of this natural product family as antineoplastic agents.^{4,7}

Biogenetically, the macrolactam core of 1 and 2 arises from sequential condensations of 4-hydroxylysine, L-alanine, and malonyl-CoA catalyzed by modules of nonribosomal peptide synthetases. In 2007,⁸ Dudler and co-workers identified the biosynthetic gene cluster responsible for the production of glidobactin from a soil bacterium belonging to *Burkholderiales*. Eight genes, named *glbA–H*, were found to be crucial for the

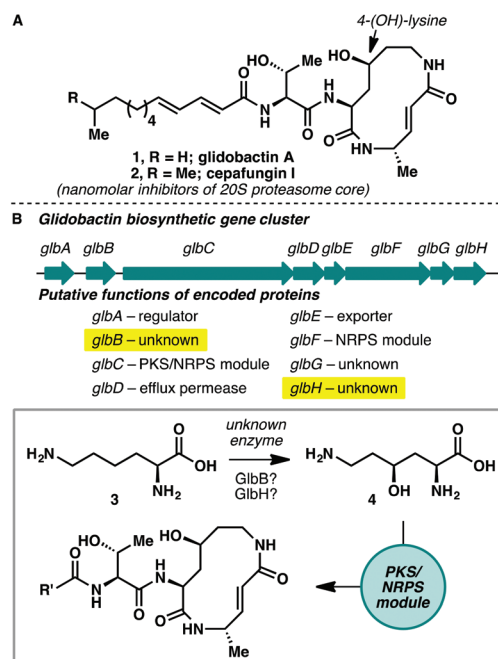


Fig. 1 (A) Structures of glidobactin A (1) and cepafungin I (2), members of the syrbactin family of natural products with potent inhibitory activity against the 20S proteasome core. (B) Biosynthetic gene cluster of glidobactin in *Polyangium brachysporum* and predicted functions of the respective genes in the formation of 1.

The Scripps Research Institute, 130 Scripps Way, Jupiter, FL 33458, USA.

E-mail: hrenata@scripps.edu

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c8ob02054j

biosynthesis of glidobactin (Fig. 1B). Two of the genes, *glbC* and *glbF*, were assigned to encode for the NRPS modules responsible for the formation of the macrolactam core structure. Of the remaining six, *glbA*, *glbD*, and *glbE* are thought to encode for proteins involved in the export and regulation of the metabolite production. GlbH shares homology with FAD-dependent nitrodioxygenases and GlbB contains a conserved domain within the DUF2257 family, hitherto a family of uncharacterized conserved bacterial proteins. Though it is likely that either of the latter two genes is responsible for the production of 4-(OH)-Lys, the exact biosynthetic origin of this motif was never elucidated on a molecular level.

In surveying the literature, several lines of evidence seem to point to the role of GlbB in the production of 4-(OH)-Lys. Firstly, deletion of *glbH* from the producer strain by Dudler led to only reduction and not total abolition of glidobactin production.⁸ A *glbB* disruption mutant, however, was found to be completely incapable of producing glidobactin. A homologous syrbactin synthetase from *Photorhabdus luminescens*, despite containing only homologs of *glbB-G*, is capable of producing glidobactin A when expressed in *Pseudomonas putida*.⁹ Finally, while the function of DUF2257 (now known as PF10014) was unknown at the time of Dudler's work, subsequent investigation by Ogawa and co-workers revealed that it belongs to the iron- and α -ketoglutarate-dependent (Fe/ α KG) superfamily of enzymes and that its members are capable of hydroxylating free-standing aliphatic amino acids.¹⁰ Indeed, a homology model of GlbB, built using a solved crystal structure of a member of PF10014 from *M. petroleiphilum* as a template,¹¹ revealed the conserved His¹-X-Asp-X_n-His² metal binding triad consisting of His188, Asp190, and His251. On the opposite side of the pocket, an arginine residue (Arg267) is likely involved in the binding of α KG via salt bridge formation. The homology model also suggests a hydrophilic pocket formed by Asp73, Arg82, Gln115, and Arg128, which likely forms key polar contacts with the amino acid substrate (Fig. S1†). Based on these observations, we postulate that GlbB might be responsible for the C4 hydroxylation of L-lysine in the biosynthesis of glidobactin. Recent results from our laboratory have argued for the general utility of Fe/ α KG enzymes as enabling tools in the production of valuable building blocks for the synthesis of complex bioactive molecules.¹² Hypothesizing that selective C4 hydroxylation of L-Lys can be leveraged for the chemoenzymatic synthesis of syrbactins and related analogues, we set out to functionally characterize GlbB.

Codon optimization and over-expression of GlbB as an N-His-tagged protein was found to provide good yield of soluble enzyme (Fig. S3†). *In vitro* assays with recombinant GlbB were next performed to test its ability to hydroxylate L-Lys (Fig. 2). In the presence of FeSO₄, α -ketoglutarate (α KG), and ascorbic acid, reaction of L-Lys with GlbB was observed to generate a new product with *m/z* = 163, corresponding to one hydroxylation event. Omission of either FeSO₄ or α KG from the reaction led to abolition of catalytic activity, confirming the dependency of the enzyme on these components for activity. Subsequent NMR characterization verified the identity of the

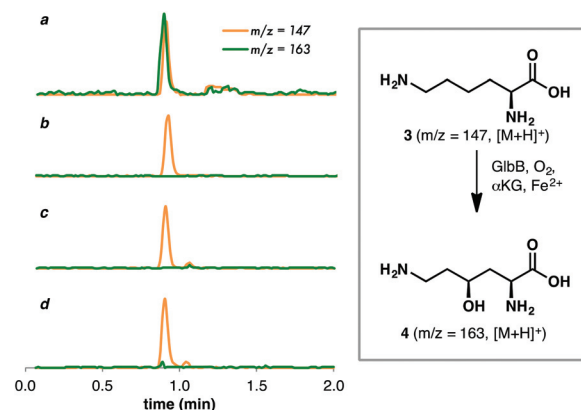


Fig. 2 LCMS profile of L-Lys hydroxylation (a) in the presence of GlbB, FeSO₄, α KG, and ascorbic acid; (b) in the absence of GlbB; (c) in the absence of FeSO₄; (d) in the absence of α KG.

reaction product as **4**, thus establishing the role of GlbB as lysine 4-hydroxylase in the biosynthesis of glidobactin.

To investigate the utility of GlbB as a hydroxylation biocatalyst, we tested its activity against a suite of free amino acid substrates. Among the 20 substrates tested (Fig. 3 and Fig. S4†), only L-Lys, L-Leu and L-Met yielded the corresponding hydroxylation products. Hydroxylation of L-Lys proceeded with complete diastereoselectivity and high total turnover number (5900 TTN). Similar to several other members of the PF10014 family, GlbB can catalyze the C4 hydroxylation of L-Leu to **5**,

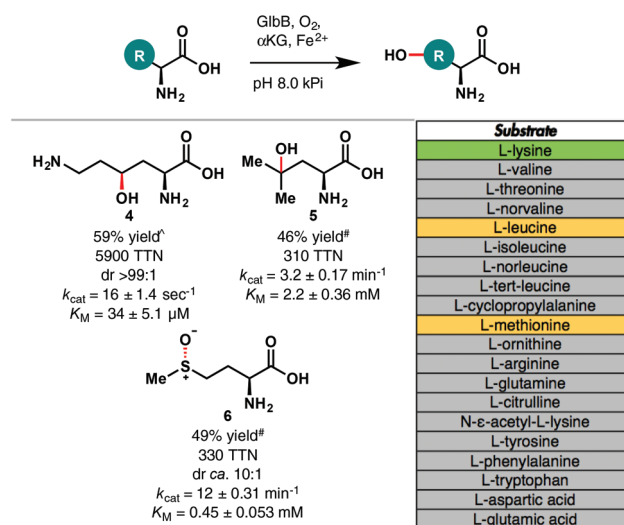
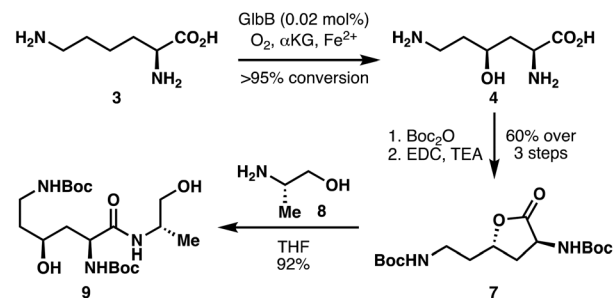


Fig. 3 Substrate scope of amino acid hydroxylation with GlbB. Conditions for total turnover number (TTN) measurement: free amino acid (20 mM, 1 equiv.), α KG (50 mM, 2.5 equiv.), L-ascorbic acid (10 mM, 0.5 equiv.), FeSO₄ (1 mM, 0.05 equiv.), GlbB (0.01–0.15 mol% enzyme loading), kPi buffer (pH = 8.0, 50 mM, 3 mL total volume), 12 h at 20 °C. Yields and TTNs were determined by NMR analysis calibrating against an internal standard of known concentration. Color legend for substrate table: green – excellent substrate, orange – moderate substrate, grey – nonsubstrate. [†]Reaction was carried out with 0.01 mol% GlbB. [‡]Reaction was carried out with 0.15 mol% GlbB.

albeit with only moderate TTN. Productive oxidation could also be observed with L-Met to provide the corresponding sulfide product **6** with 330 TTN and *ca.* 10 : 1 dr. Additionally, steady-state kinetic parameters were obtained for hydroxylation of L-Lys, L-Leu, and L-Met (Fig. S5†). Consistent with the observed TTNs, the catalytic efficiency of L-Lys hydroxylation of L-Lys was found to be a few orders of magnitude higher than that of L-Leu or L-Met, arising from significantly higher k_{cat} and much lower K_{M} . With the exception of PluL—an Fe/αKG involved in the biosynthesis of carbamoylpolyoxamic acid¹³—previously characterized members of PF10014 have mostly been limited to Fe/αKGs accepting only amino acid substrates with aliphatic side-chains.¹⁰ To the best of our knowledge, GlbB is the first member of this family capable of acting on both polar and aliphatic amino acids. Prior to our work, the two Fe/αKG lysine hydroxylases characterized thus far, KDO1 and KDO2,¹⁴ are classified as clavamate-synthase-like Fe/αKG dioxygenases (IPR014503 family), with minimal sequence similarity to GlbB. In fact, despite catalyzing the hydroxylation of the same carbon of L-Lys, KDO2 and GlbB share only 11% sequence identity. The identification of lysine hydroxylation activity within both PF10014 and IPR014503 families suggests convergent evolution of this function from highly dissimilar ancestral enzymes.

In agreement with GlbB's unique substrate specificity, a global survey of the PF10014 family, generated using the sequence similarity network (SSN) approach, shows that GlbB belongs to a distinct cluster separate from previously characterized members of PF10014 (Fig. 4). Notably, this cluster also contains an Fe/αKG encoded by *plu1881* (Uniprot accession code: Q7N5R2, 48% sequence identity with GlbB) from the glidobactin synthetase of *P. luminescens*.⁹ Assuming that enzymes within this cluster exhibit similar substrate preference, our analysis would predict that Plu1881 is also capable of catalyz-



Scheme 1 Application of L-Lys hydroxylation with GlbB in the synthesis of a key dipeptide fragment (**9**) of glidobactin.

ing C4 hydroxylation of L-Lys. More importantly, the SSN of PF10014 reveals the existence of several clusters with no functionally characterized members. For example, we identify a cluster containing HrmJ, an Fe/αKG that had been implicated in the formation of 3-(*trans*-2'-nitrocyclopropyl)alanine from L-Lys in the biosynthesis of hormaomycin.¹⁵ We anticipate that functional characterization of enzymes belonging to these lesser-known clusters will expand the breadth of chemical reaction space spanned by the Fe/αKG dioxygenase superfamily.

With an eye towards a chemoenzymatic total synthesis of the syrbactins, we next assessed the utility of GlbB in preparative scale hydroxylation of L-Lys (Scheme 1). At 0.6 mmol scale, full conversion of **3** to **4** could be obtained by using just 0.02 mol% of purified GlbB. Similar to other C4 and C5 hydroxylated amino acids,^{12,14} **4** was prone to undergo intramolecular cyclization to gamma lactone **7** during protection of the free amines. Further stirring of the crude material from the amine protection in the presence of EDC allowed the lactonization to proceed to full conversion, providing **7** in 60% isolated yield. Gratifyingly, **7** was found to undergo facile amide bond formation with L-alaninol (**8**) at room temperature to afford dipeptide **9** in 92% yield.¹⁶ This sequence illustrates the synthetic utility of GlbB as a key dipeptide fragment of glidobactin could be obtained in just four simple steps.

Conclusions

Through a series of *in vitro* experiments, we have functionally characterized a new Fe/αKG dioxygenase that is capable of catalyzing selective C4 hydroxylation of L-Lys. These studies not only shed light on the origin of the 4-(OH)-Lys motif in glidobactin, but also lay the groundwork for the development of a chemoenzymatic strategy to access the syrbactin family of natural products and related structural analogues.

Conflicts of interest

There are no conflicts to declare.

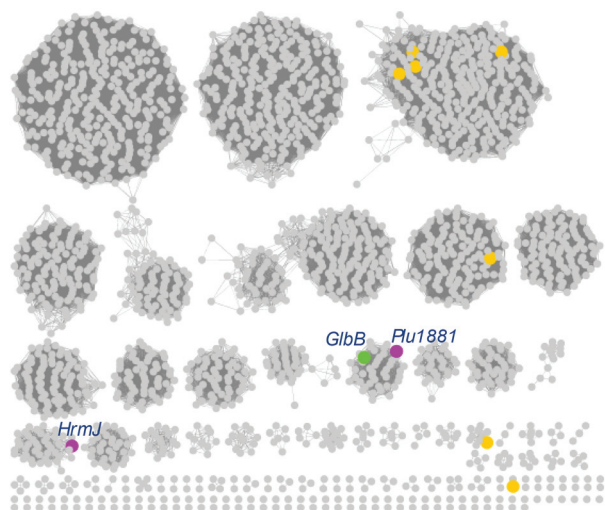


Fig. 4 Sequence similarity network of PF10014 (100% identical sequences are conflated, alignment score 60). Orange nodes represent previously characterized Fe/αKG belonging to this family.

Acknowledgements

Financial support for this work was generously provided by The Scripps Research Institute and the National Institutes of Health (Grant R35 GM128895). We thank the Shen lab, the Kodadek lab, and the Roush lab for generous access to their instrumentations.

Notes and references

- 1 Y. Terui, J. Nishikawa, H. Hinoo, T. Kato and J. Shoji, *J. Antibiot.*, 1990, **43**, 788; U. Wäspi, P. Hassa, A. Staempfli, L.-P. Molleyres, T. Winkler and R. Dudler, *Microbiol. Res.*, 1999, **154**, 89.
- 2 M. Groll, B. Schellenberg, A. S. Bachmann, C. R. Archer, R. Huber, T. K. Powell, S. Lindow, M. Kaiser and R. Dudler, *Nature*, 2008, **452**, 755.
- 3 M. L. Stein, P. Beck, M. Kaiser, R. Dudler, C. F. W. Becker and M. Groll, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 18367.
- 4 J. Clerc, M. Groll, D. J. Illich, A. S. Bachmann, R. Huber, B. Schellenberg, R. Dudler and M. Kaiser, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 6507.
- 5 A. Rentsch, D. Landsberg, T. Broadmann, L. Bülow, A.-K. Girbig and M. Kalesse, *Angew. Chem., Int. Ed.*, 2013, **52**, 5450; J. S. Schneekloth Jr. and C. M. Crews, *Curr. Drug Targets*, 2011, **12**, 1581; S. Tsukamoto and H. Yokosawa, *Planta Med.*, 2010, **76**, 1064.
- 6 P. M. Cromm and C. M. Crews, *ACS Cent. Sci.*, 2017, **3**, 830.
- 7 U. Schmidt, A. Kleefeldt and R. Mangold, *J. Chem. Soc., Chem. Commun.*, 1992, 1687; T. Chiba, H. Hosono, K. Nakagawa, M. Asaka, H. Takeda, A. Matsuda and S. Ichikawa, *Angew. Chem., Int. Ed.*, 2014, **53**, 4836; M. C. Pirrung, G. Biswas and T. R. Ibarra-Rivera, *Org. Lett.*, 2010, **12**, 2402; S. Kitahata, T. Chiba, T. Yoshida, M. Ri, S. Iida, A. Matsuda and S. Ichikawa, *Org. Lett.*, 2016, **18**, 2312; S. Kitahata, F. Yakushiji and S. Ichikawa, *Chem. Sci.*, 2017, **8**, 6959; J. Opoku-Ansah, T. R. Ibarra-Rivera, M. C. Pirrung and A. S. Bachmann, *Pharm. Biol.*, 2012, **50**, 25.
- 8 B. Schellenberg, L. Bigler and R. Dudler, *Environ. Microbiol.*, 2007, **9**, 1640.
- 9 A. Dudnik, L. Bigler and R. Dudler, *Microbiol. Res.*, 2013, **168**, 73.
- 10 T. Kodera, S. V. Smirnov, N. N. Samsonova, Y. I. Kozlov, R. Koyama, M. Hibi, J. Ogawa, K. Yokozeki and S. Shimizu, *Biochem. Biophys. Res. Commun.*, 2009, **390**, 506; M. Hibi and J. Ogawa, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 3869.
- 11 Q. Xu, *et al.*, *Proteins*, 2014, **82**, 164.
- 12 C. R. Zwick III and H. Renata, *J. Am. Chem. Soc.*, 2018, **140**, 1165; X. Zhang, E. King-Smith and H. Renata, *Angew. Chem., Int. Ed.*, 2018, **57**, 5037.
- 13 J. Qi, D. Wan, H. Ma, Y. Liu, R. Gong, X. Qu, Y. Sun, Z. Deng and W. Chen, *Cell Chem. Biol.*, 2016, **23**, 935.
- 14 D. Baud, *et al.*, *ChemCatChem*, 2014, **6**, 3012.
- 15 I. Höfer, M. Crüsemann, M. Radzom, B. Geers, D. Flachshaar, X. Cai, A. Zeeck and J. Piel, *Chem. Biol.*, 2011, **18**, 381.
- 16 J. Marin, J.-P. Briand and G. Guichard, *Eur. J. Org. Chem.*, 2008, 1005.