

Biosynthesis

Genome Mining of Oxidation Modules in *trans*-Acyltransferase Polyketide Synthases Reveals a Culturable Source for LobatamidesReiko Ueoka⁺, Roy A. Meoded⁺, Alejandro Gran-Scheuch, Agneya Bhushan, Marco W. Fraaije, and Jörn Piel*

Abstract: Bacterial *trans*-acyltransferase polyketide synthases (*trans*-AT PKSs) are multimodular megaenzymes that biosynthesize many bioactive natural products. They contain a remarkable range of domains and module types that introduce different substituents into growing polyketide chains. As one such modification, we recently reported Baeyer–Villiger-type oxygen insertion into nascent polyketide backbones, thereby generating malonyl thioester intermediates. In this work, genome mining focusing on architecturally diverse oxidation modules in *trans*-AT PKSs led us to the culturable plant symbiont *Gyneuella sunshinyii*, which harbors two distinct modules in one orphan PKS. The PKS product was revealed to be lobatamide A, a potent cytotoxin previously only known from a marine tunicate. Biochemical studies show that one module generates glycolyl thioester intermediates, while the other is proposed to be involved in oxime formation. The data suggest varied roles of oxygenation modules in the biosynthesis of polyketide scaffolds and support the importance of *trans*-AT PKSs in the specialized metabolism of symbiotic bacteria.

Bacterial complex polyketides belong to the most important natural product classes of therapeutic value.^[1] Two distinct families of multimodular polyketide synthases (PKSs), termed *cis*- and *trans*-acyltransferase (*trans*-AT) PKSs, generate most of these compounds. The textbook biosynthetic model is represented by the *cis*-AT PKS family, such as the erythromycin PKS.^[1,2] These enzymes are usually composed of a limited set of functionally different biosynthetic multi-domain modules that elongate and modify intermediates, and give rise to contiguous carbon chains carrying keto, hydroxy, double-bond, and carbon branch modifications. In contrast, the highly complex *trans*-AT PKSs can accommodate a re-

markable diversity of modules (> 150 module architectures are currently known^[3]) and employ integrated domains as well as free-standing, *trans*-acting enzymes with often poorly understood functions.^[4] The large functional range of *trans*-AT PKSs suggests high biosynthetic diversity outside the scope of canonical polyketides. Another phenomenon of *trans*-AT PKSs is their dominant role in the specialized metabolism of non-actinomycete symbiotic bacteria, with examples reported from fungi,^[5] insects,^[6] marine invertebrates,^[7] and plants.^[8] To fully access this biosynthetic potential, insights into the function of *trans*-AT PKS assembly lines and their components are needed to improve our ability to predict their products,^[1] identify alternative producers to invertebrate sources,^[9] and provide novel tools for biosynthetic engineering.



In previous work, we identified the insertion of oxygen into growing polyketide chains as a non-canonical reaction by which *trans*-AT PKSs diversify product skeletons.^[9] As shown for the *trans*-acting flavin adenine dinucleotide (FAD)-dependent oxygenase OocK from the oocydin pathway, the enzyme performs a Baeyer–Villiger (BV) oxidation on a β -keto thioester intermediate to generate a malonyl derivative (Figure 1 A) that is then further elongated. Preservation or hydrolysis of the new ester moiety gives rise to oxygen atoms within polyketide backbones (oocydin B and haterumalides), carboxylate pseudostarters (oocydin A), or terminal alcohols (introduced by the OocK homologues in the pederin^[6] and toblerol pathway,^[8] Figure 1 B). Mid-chain oxygen atoms also occur in other polyketides (Figure 1 C).^[10] Although these moieties could principally result from various processes,^[11] the presence of OocK homologs in diverse orphan PKS biosynthetic gene clusters (BGCs; Figures S1–S4) suggests that nature employs oxygen insertion more widely to construct unusual polyketide scaffolds. In this work, we applied a genome-mining strategy to investigate the wider scope of flavoprotein monooxygenases acting during polyketide elongation. The data revealed an architecturally distinct oxygen insertion module that generates a glycolyl intermediate, as well as evidence for a third module type involved in oxime formation. Both new modules are used in the biosynthesis of lobatamides, potent cytotoxins that were previously only known from marine tunicates, but identified here from a culturable plant symbiont.

We initiated our work by identifying OocK homologues and other PKS-associated flavoprotein monooxygenases encoded in orphan BGCs using GenBank, genome neighborhood,^[12] phylogenetic, and manual analyses (status November 2019; Figures S1–S4 in the Supporting Information). These revealed 69 candidate *trans*-AT PKS systems harboring such

[*] Dr. R. Ueoka,^[†] Dr. R. A. Meoded,^[†] Dr. A. Bhushan, Prof. Dr. J. Piel
Institute of Microbiology, ETH Zurich
Vladimir-Prelog-Weg 4, 8093 Zurich (Switzerland)
E-mail: jpiel@ethz.ch

A. Gran-Scheuch, Prof. Dr. M. W. Fraaije
Molecular Enzymology Group, University of Groningen
Nijenborgh 4, 9747AG Groningen (The Netherlands)
A. Gran-Scheuch
Department of Chemical and Bioprocesses Engineering
Pontificia Universidad Católica de Chile
Avenida Vicuña Mackenna 4860, 7820436 Santiago (Chile)

[†] These authors contributed equally to this work.

 Supporting information (including experimental details) and the
 ORCID identification number(s) for the author(s) of this article can
be found under:
<https://doi.org/10.1002/anie.201916005>.

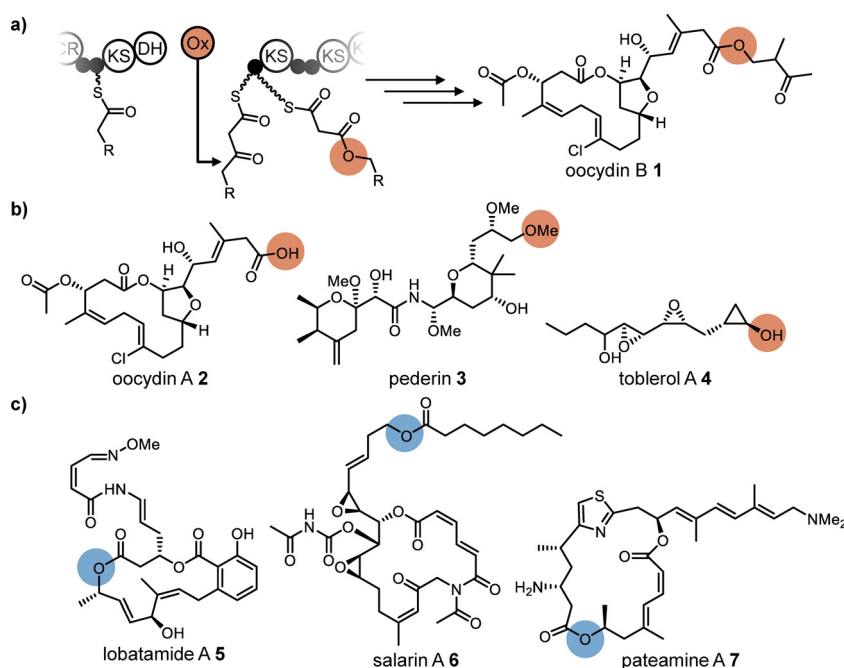


Figure 1. Oxygen incorporation into polyketide backbones. Moieties derived from oxygen insertion are highlighted in orange. Oxygen atoms of unknown biosynthetic origin are highlighted in blue. a) Mechanism for oxygen insertion into growing polyketide backbones, exemplified by OocK, a trans-acting BV monooxygenase from the oocidin (oocidin B, 1) pathway. The enzyme converts a β -ketothioester substrate into a malonyl derivative. b) Oocidin A (2), pederin (3), and toblerol A (4) with carboxylate- and alcohol-type termini arising from oxygen insertion and ester cleavage. c) Biosynthetically unassigned polyketides harboring mid-chain oxygens: lobatamide A (5) from a tunicate, and salarin A (6) and pateamine A (7) from marine sponges.

enzymes (Figures S1–S4). The phylogram contained a large clade containing the functionally related OocK, PedG, and TobD, as well as homologues from uncharacterized PKSs. Another functionally assigned clade bears module-integrated oxygenase (Ox) domains for which the variant from *Burkholderia* sp. FERM BP-3421 was shown to introduce an epoxide unit in spliceostatin (Figure S1).^[8] In addition, other non-OocK clades mostly belong to Ox domains integrated in various module types, thus suggesting further biosynthetic diversity. Focusing on uncharacterized oxygenases (Figure S1), an orphan *trans*-AT system (termed *lbm* PKS, Figure 2A) in the Gram-negative bacterium *Gyvuella sunshinyi* (NZ_CP007142) appeared an intriguing candidate, since the PKS contains oxygenase modules from two different unassigned clades (Figure S1, Table S1). *G. sunshinyi* is an unusual halophilic root-associated plant symbiont^[13] that was recently recognized as a talented producer of diverse natural products.^[14] Its genome contains six *trans*-AT PKS BGCs, the highest known number for any organism. The two integrated Ox domains are located in the PKS proteins LmbA and LmbC, and will be referred to as LmbA-Ox and LmbC-Ox, respectively. Additional features in these modules are a methyltransferase (MT) in the LmbC-Ox module, which suggests α -C-methylation, and a predicted non-elongating KS (KS⁰) in the LmbA-Ox module, which suggests that a moiety introduced by the upstream module is further modified. This upstream module is located at the N terminus of LmbA and

resembles an NRPS loading module with predicted^[15] glycine specificity. A C-terminal TE domain is encoded on the PKS gene *lbmE* at the downstream end, which overall suggests collinearity between the PKS gene order and biosynthetic events.

As guidance for the targeted isolation of the *lbm* polyketide, we analyzed the PKS using the recently developed automated prediction tool TransAFor.^[3] This web application suggests chemical structures for *trans*-AT PKSs from phylogenetically inferred^[15] KS substrates. In agreement with the unusual PKS architecture, the TransAFor output 8 (Figure 2B) contained extended regions of low confidence. After removal of a methyl group from the five-bonded C18, the structure was refined by closer inspection of the modular architecture. Since TransAFor can predict intermediates only for modules that have a downstream KS, no elongation was predicted for the terminal module. A terminal C2 unit was therefore added based on *cis*-AT PKS rules. To account for the N-terminal NRPS module of LbmA, we suspected glycine for biosynthetic initiation. This replacement also removed an exomethylene group in the TransAFor structure that was unlikely because polyketide β -branching compo-

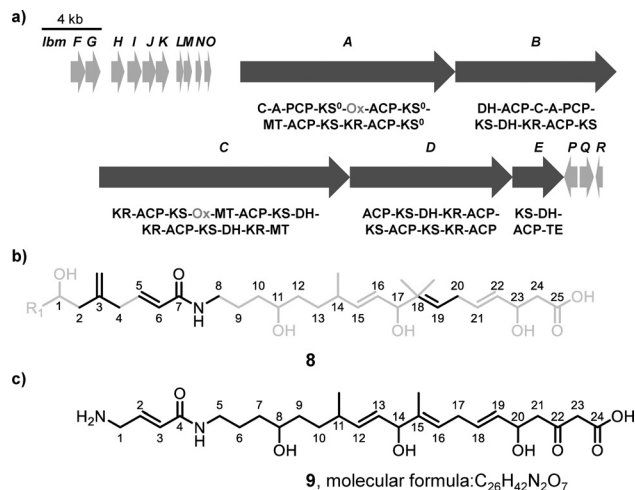


Figure 2. The *lbm* BGC and predicted polyketide structures. a) BGC architecture with core PKS and accessory genes marked in dark and light grey, respectively. The domain architecture of the core PKSs is shown using the following abbreviations: A = adenylation domain; C = condensation domain; DH = dehydratase; ER = enoylreductase; KR = ketoreductase; KS = ketosynthase; KS⁰ = non-elongating KS; MT = methyltransferase; Ox = oxygenase; TE = thioesterase. Ox domains are shown in grey. b) TransAFor-predicted structure 8 of the *lbm* product. High- and low-confidence regions are shown in black and grey, respectively. c) Manually refined structure 9 used to find the natural product.

nents^[16] were not found in the BGC. These modifications resulted in the hypothetical structure **9** (Figure 2C) as a basis for the analytical work.

To search for the predicted compound, *G. sunshinyi* was cultivated, and the extract was analyzed by ultra-high-performance liquid chromatography/high-resolution mass spectrometry (UHPLC–HRMS). Manual inspection of metabolite-related MS features and the molecular formulae suggested from the high-resolution masses showed a candidate ion peak at m/z 513.2230 $[M+H]^+$ (Figure S5). This corresponds to a molecular formula of $C_{27}H_{32}N_2O_8$ (calc. 513.2231). This formula was closest to that of the predicted structure ($C_{26}H_{42}N_2O_7$) and appeared a good candidate for the *lbm* product. No other orphan PKSs in the genome were predicted to account for a similar chemical formula that contains two nitrogen atoms (i.e., PKSs that contain two NRPS modules).^[14] To isolate the compound, bacterial pellets from a 2 L culture of *G. sunshinyi* culture were extracted with acetone. MS-guided fractionation using reversed phase-high performance liquid chromatography (RP-HPLC) provided pure compound **5**. The ¹H NMR and HSQC data showed the presence of an aliphatic methyl coupling with a vicinal proton, one methyl connected to an sp² carbon, one methoxy group, three oxymethines, and 11 methines connected to sp² carbons (Figure S6). Analysis of 2D NMR data including COSY, HMBC, and HSQC revealed the structure as that of lobatamide A (Figure 3A, Figures S7–S9). Lobatamides are potent vacuolar (H⁺)-ATPase inhibitors that interfere with tumor metastasis.^[17] Since they were previously known only from a marine invertebrate, the tunicate *Aplidium lobatum*,^[18] *G. sunshinyi* represents the first culturable source for these compounds. Lobatamides belong to a diverse group of benzolactone (H⁺)-ATPase inhibitors (Figure 3A) with as-yet unknown PKSs and isolated from a remarkable range of organisms, including a sponge (salicylihalamides),^[19] a fungus (CJ-12,950),^[20] myxobacteria (apicularens),^[21] and a *Pseudomonas* strain (oximidines).^[22] Their related structures and the identification of bacterial sources suggest that all of these compounds are of prokaryotic origin. For an expansion of this family, see a study on necroximes that was conducted concurrently with this work.^[23]

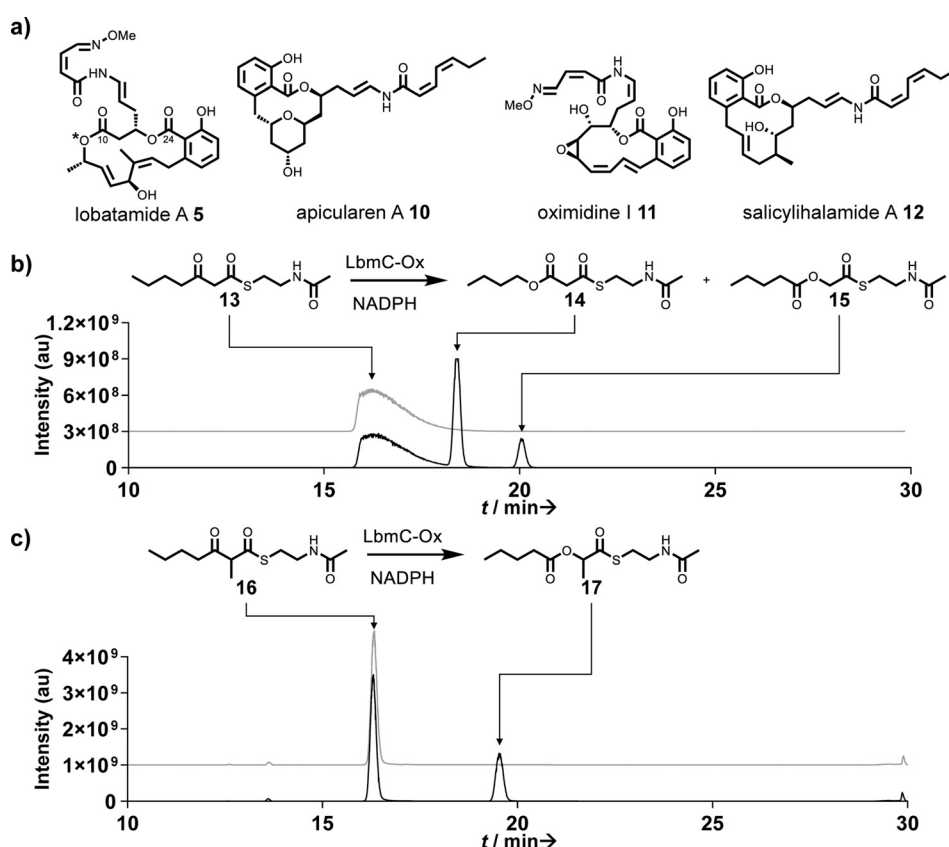


Figure 3. Selected benzolactone enamide polyketides and enzymatic assays with LbmC-Ox. a) Benzolactone enamides. The moiety derived from oxygen insertion is highlighted with an asterisk. b, c) UHPLC–HRMS data showing the extracted ion chromatograms (EIC) of assay mixtures, including a boiled-enzyme negative control (upper/grey) and a test reaction using all components (lower/black). b) EIC for **13** and **13** + [¹⁶O] (calc. for $[M+H]^+$ as 246.1158, and 262.1108, respectively). For the mass spectrum of the new product **15** see Figure S11. c) EIC for **16** and **16** + [¹⁶O] (calc. for $[M+H]^+$ as 260.1315 and 276.1264, respectively). Mass spectrum of product **17** is shown in Figure S19.

Structure **5** features an internal ester that is reversed in comparison to oocydin, pederin, and toblerol biosynthesis. To investigate the role of the LbmC-Ox module in its formation, we cloned the region encoding the Ox domain into a variant of the pET28a expression vector (see the Supporting Information) for expression in *E. coli* as an N-terminally His₆-tagged protein (Figure S10). Purified LbmC-Ox was obtained as a soluble yellow protein, thus indicating bound FAD. Since the LbmC-Ox module contained an additional MT domain, we suspected that the Ox domain accepts either a β -keto-thioester or an α -methyl- β -keto-thioester, depending on whether LbmC-Ox acts before or after C-methylation. We initially assayed the enzyme with the unmethylated thioester **13** as a simplified surrogate of an acyl carrier protein (ACP)-bound 4'-phosphopantetheinyl intermediate. Assay mixtures with test substrate, LbmC-Ox, and NADPH were incubated for 90 minutes at room temperature and then extracted with ethyl acetate. UHPLC–HRMS analysis suggested two new products in the assay mixture, both with a mass difference of +16 Da relative to the substrate (Figure 3B). Isolation and NMR-based structure elucidation of the products (Figures S12–17, Table S4) revealed that the two isomers carry the inserted oxygen either at the γ (Figure 3B, product **14**) or

the β position (product **15**). The less abundant thioester **15** is noteworthy because it exhibits a reversed ester moiety compared to the OocK/PedG/TobD products and corresponds to the topology present in lobatamides. Suspecting that the major isomer **14** was an aberrant byproduct due to the choice of the surrogate substrate, we also synthesized thioester **16**, which harbors the predicted methyl group at the α position (Figure 3C, Figure S18). Repetition of the enzyme assay with this new test substrate produced one product peak (Figure 3C). Isolation and structure elucidation identified it as thioester **17** (Figure 3C, Figure S19–24, Table S5). The result from the biochemical study matches the internal ester topology at C-10 of lobatamides, that is, facing the same direction as the C-24 ester created by macrolactonization. In lieu of knock-out studies that were so far unsuccessful or of heterologous whole-cluster expression, the unusual function of LbmC-Ox and its location within the PKS reasonably link the *lbm* BGC to lobatamide.

The results suggest an overall biosynthetic model for lobatamides as shown in Figure 4. A glycine starter would be the source of the oxime moiety, and oxygen insertion and macrolactonization generate the ester groups at C-10 and C-24, respectively. The salicylate group generated by the terminal four modules is also a feature of all other members of the benzolactone family, which currently lack known BGCs. However, similar tetramodular series that generate benzene moieties are also known from the otherwise unrelated psymberin and legiolulin PKSs.^[24]

Comparing the PKS-associated enzymes OocK and LbmC-Ox to a wider range of previously characterized class B flavoprotein monooxygenases (Figure S25), BV oxidation activity appears more prevalent in the LbmC-Ox than the OocK clade.^[25] To further explore their biocatalytic potential, we tested 37 substrates that do not resemble polyketide intermediates and harbor diverse functional

groups (Figures S26–S27, Table S6). When comparing the free-standing OocK with the excised LbmC-Ox, the latter exhibited greater thermostability and a broad substrate scope, accepting 12 of the 37 test compounds (**18–26**, Figure 5). Both enzymes accept NADPH as hydride donor, while OocK also displays activity with NADH (Table S7). These results suggest that the two enzymes could be used as biocatalysts for various oxidations, in addition to showing potential as PKS engineering tools to introduce non-standard moieties into polyketides.

In conclusion, our data support the existence of at least two further module types in *trans*-AT PKSs that permit oxidative modifications of polyketide core structures. In addition to the topologically related oxygen insertions catalyzed by modules utilizing *trans*-acting OocK/PedG/TobD-type enzymes (clade I modules, Figure S1), we identified a distinct, module-integrated Ox domain that installs esters of the opposite orientation (clade II modules, Figure S1). It is unknown whether all members of one oxygenase clade generate the same ester topology. However, in line with

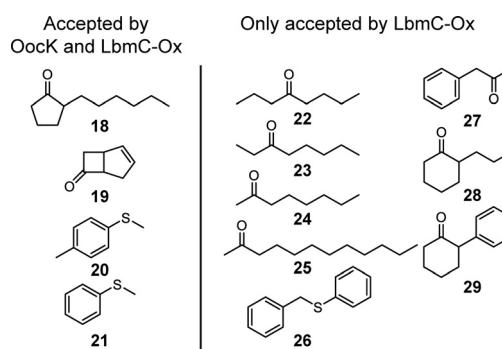


Figure 5. Structures of compounds that deviate from polyketide intermediate but were accepted by LbmC-Ox and OocK or by LbmC-Ox alone. For products, see Figure S26.

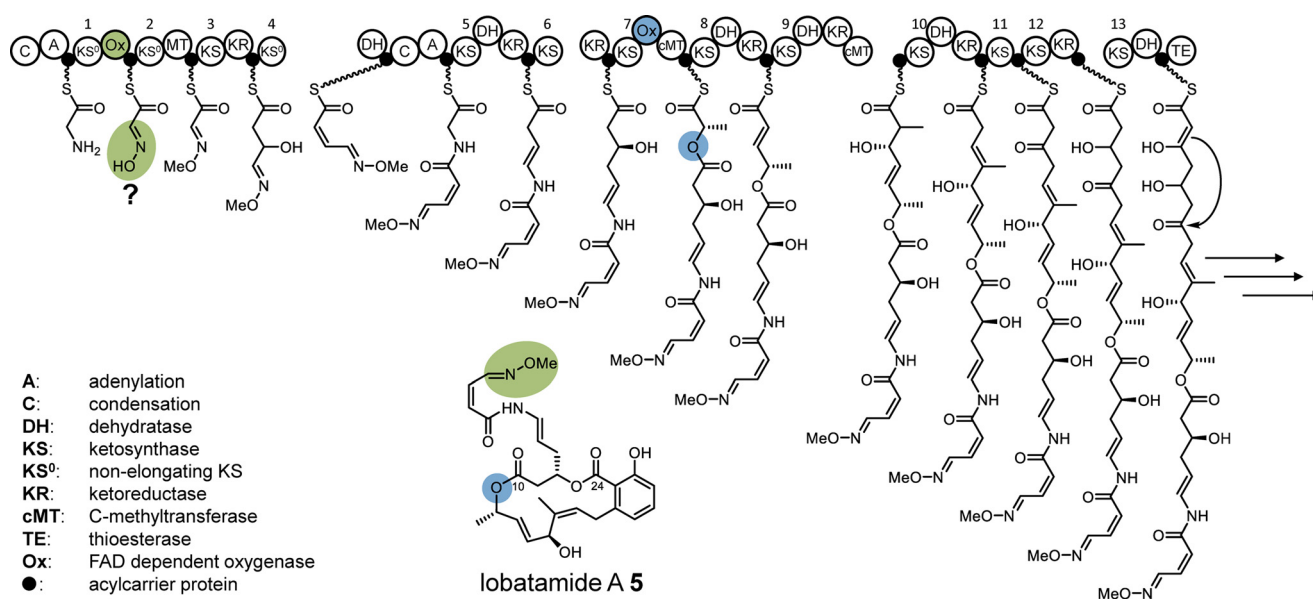


Figure 4. Biosynthetic model for lobatamide A. Ox domains and the corresponding polyketide moieties (proposed for the oxime) are highlighted in color. Consecutive KS numbers are shown above the domains.

a potentially predictive uniform product pattern for clade II, an uncharacterized oxygenase from the mandelalide *trans*-AT PKS^[7c] that is analogous to LmbC-Ox has been suggested to insert oxygen in the same orientation as for lobatamides.^[7c] A second type of oxygenase-containing module newly identified in this work is not involved in oxygen insertion but seems to play a role in glycine oxidation to generate the unusual oxime moiety of lobatamides. This hypothesis is currently based on collinearity logic and needs to be experimentally tested. These examples demonstrate the staggering functional diversity of *trans*-AT PKSs, which not only introduce substituents but also modify the carbon backbone itself. Our results also exemplify how genome mining can pave the way to finding microbial sources for bioactive compounds that had previously been described only from marine invertebrates. The production of the tunicate metabolite lobatamide A by *G. sunshinyii* further validates this plant symbiont as a rich source of bioactive specialized metabolites.

Acknowledgements

We thank Young Ryun Chung and Dmitri Mavrodi for discussions on the biology of *G. sunshinyii*. This project was funded by the SNSF (grant IDs 205321_165659 and 205320_185077), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 742739).

Conflict of interest

The authors declare no conflict of interest.

Keywords: bacterial natural products · biosynthesis · marine natural products · polyketides

How to cite: *Angew. Chem. Int. Ed.* **2020**, *59*, 7761–7765
Angew. Chem. **2020**, *132*, 7835–7839

- [1] C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716; *Angew. Chem.* **2009**, *121*, 4782–4811.
- [2] C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, *Annu. Rev. Biochem.* **2007**, *76*, 195–221.
- [3] E. J. Helfrich, R. Ueoka, A. Dolev, M. Rust, R. A. Meoded, A. Bhushan, G. Califano, R. Costa, M. Gugger, C. Steinbeck, *Nat. Chem. Biol.* **2019**, *15*, 813–821.
- [4] a) J. Piel, *Nat. Prod. Rep.* **2010**, *27*, 996–1047; b) E. J. Helfrich, J. Piel, *Nat. Prod. Rep.* **2016**, *33*, 231–316.
- [5] L. P. Partida-Martinez, C. Hertweck, *ChemBioChem* **2007**, *8*, 41–45.
- [6] J. Piel, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14002–14007.
- [7] a) R. Ueoka, A. R. Uria, S. Reiter, T. Mori, P. Karbaum, E. E. Peters, E. J. Helfrich, B. I. Morinaka, M. Gugger, H. Takeyama, S. Matsunaga, J. Piel, *Nat. Chem. Biol.* **2015**, *11*, 705–712; b) M. Hildebrand, L. E. Waggoner, H. B. Liu, S. Sudek, S. Allen, C. Anderson, D. H. Sherman, M. Haygood, *Chem. Biol.* **2004**, *11*, 1543–1552; c) J. Lopera, I. J. Miller, K. L. McPhail, J. C. Kwan, *mSystems* **2017**, *2*, e00096-00017.
- [8] R. Ueoka, M. Bortfeld-Miller, B. I. Morinaka, J. A. Vorholt, J. Piel, *Angew. Chem. Int. Ed.* **2018**, *57*, 977–981; *Angew. Chem.* **2018**, *130*, 989–993.
- [9] R. A. Meoded, R. Ueoka, E. J. Helfrich, K. Jensen, N. Magnus, B. Piechulla, J. Piel, *Angew. Chem. Int. Ed.* **2018**, *57*, 11644–11648; *Angew. Chem.* **2018**, *130*, 11818–11822.
- [10] a) A. Bishara, A. Rudi, M. Aknin, D. Neumann, N. Ben-Califa, Y. Kashman, *Org. Lett.* **2008**, *10*, 153–156; b) P. T. Northcote, J. W. Blunt, M. H. Munro, *Tetrahedron Lett.* **1991**, *32*, 6411–6414.
- [11] a) Z. D. Dunn, W. J. Wever, N. J. Economou, A. A. Bowers, B. Li, *Angew. Chem. Int. Ed.* **2015**, *54*, 5137–5141; *Angew. Chem.* **2015**, *127*, 5226–5230; b) J. B. Biggins, M. A. Ternei, S. F. Brady, *J. Am. Chem. Soc.* **2012**, *134*, 13192–13195; c) J. Franke, K. Ishida, C. Hertweck, *Angew. Chem. Int. Ed.* **2012**, *51*, 11611–11615; *Angew. Chem.* **2012**, *124*, 11779–11783.
- [12] R. Zallot, N. Oberg, J. A. Gerlt, *Biochemistry* **2019**, *58*, 4169–4182.
- [13] E. J. Chung, J. A. Park, C. O. Jeon, Y. R. Chung, *Int. J. Syst. Evol. Microbiol.* **2015**, *65*, 1038–1043.
- [14] R. Ueoka, A. Bhushan, S. I. Probst, W. M. Bray, R. S. Lokey, R. G. Linington, J. Piel, *Angew. Chem. Int. Ed.* **2018**, *57*, 14519–14523; *Angew. Chem.* **2018**, *130*, 14727–14731.
- [15] T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck, J. Piel, *Nat. Biotechnol.* **2008**, *26*, 225–233.
- [16] C. T. Calderone, W. E. Kowtoniuk, N. L. Kelleher, C. T. Walsh, P. C. Dorrestein, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8977–8982.
- [17] M. Pérez-Sayáns, J. M. Somoza-Martín, F. Barros-Angueira, J. M. G. Rey, A. García-García, *Cancer Treat. Rev.* **2009**, *35*, 707–713.
- [18] T. C. McKee, D. L. Galinis, L. K. Pannell, J. H. Cardellina, J. Laakso, C. M. Ireland, L. Murray, R. J. Capon, M. R. Boyd, *J. Org. Chem.* **1998**, *63*, 7805–7810.
- [19] X.-S. Xie, D. Padron, X. Liao, J. Wang, M. G. Roth, J. K. De Brabander, *J. Biol. Chem.* **2004**, *279*, 19755–19763.
- [20] K. A. Dekker, R. J. Aiello, H. Hirai, T. Inagaki, T. Sakakibara, Y. Suzuki, J. F. Thompson, Y. Yamauchi, N. Kojima, *J. Antibiot.* **1998**, *51*, 14–20.
- [21] B. Kunze, R. Jansen, F. Sasse, G. Höfle, H. Reichenbach, *J. Antibiot.* **1998**, *51*, 1075–1080.
- [22] J. W. Kim, K. Shin-ya, K. Furihata, Y. Hayakawa, H. Seto, *J. Org. Chem.* **1999**, *64*, 153–155.
- [23] S. P. Niehs, B. Dose, S. Richter, S. J. Pidot, H. Dahse, T. P. Stinear, C. Hertweck, *Angew. Chem. Int. Ed.* **2020**, <https://doi.org/10.1002/anie.201916007>; *Angew. Chem.* **2020**, <https://doi.org/10.1002/ange.201916007>.
- [24] a) K. M. Fisch, C. Gurgui, N. Heycke, S. A. van der Sar, S. A. Anderson, V. L. Webb, S. Taudien, M. Platzer, B. K. Rubio, S. J. Robinson, P. Crews, J. Piel, *Nat. Chem. Biol.* **2009**, *5*, 494–501; b) T. Ahrendt, M. Miltenberger, I. Haneburger, F. Kirchner, M. Kronenwerth, A. O. Brachmann, H. Hilbi, H. B. Bode, *ChemBioChem* **2013**, *14*, 1415–1418.
- [25] A. Riebel, M. J. Fink, M. D. Mihovilovic, M. W. Fraaije, *ChemCatChem* **2014**, *6*, 1112–1117.

Manuscript received: December 13, 2019

Accepted manuscript online: February 10, 2020

Version of record online: March 19, 2020