Biosynthesis

An Iterative Module in the Azalomycin F Polyketide Synthase Contains a Switchable Enoylreductase Domain

Wei Xu[+], Guifa Zhai[+], Yuanzhen Liu, Yuan Li, Yanrong Shi, Kui Hong, Hui Hong, Peter F. Leadlay, Zixin Deng, and Yuhui Sun*

Abstract: *Detailed analysis of the modular Type I polyketide synthase (PKS) involved in the biosynthesis of the margino-lactone azalomycin F in mangrove Streptomyces sp. 211726 has shown that only nineteen extension modules are required to accomplish twenty cycles of polyketide chain elongation. Analysis of the products of a PKS mutant specifically inactivated in the dehydratase domain of extension-module 1 showed that this module catalyzes two successive elongations with different outcomes. Strikingly, the enoylreductase domain of this module can apparently be "toggled" off and on : it functions in only the second of these two cycles. This novel mechanism expands our understanding of PKS assembly-line catalysis and may explain examples of apparent non-colinear-ity in other modular PKS systems.*

Modular Type I polyketide synthases (PKSs) catalyze the biosynthesis of numerous pharmacologically relevant natural products that exhibit antibacterial, antifungal, anthelmintic, antitumor, or immunosuppressive activities.[1–4] These giant multimodular enzymes form a processive assembly-line to produce the polyketide backbone using (alkyl)malonyl-CoA esters as the source of extender units. Each module consists of a β-ketoacyl synthase (KS) domain for condensing the incoming unit onto the growing polyketide chain, an acyl-transferase (AT) domain for loading extension units, and an acyl carrier protein (ACP) domain for retention of the growing polyketide chain on the PKS. In addition to these conserved domains, β-ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains are optional to achieve varying degrees of reduction. Finally, the full-length chain is released, for example by a cyclase/thioesterase

(TE).[5] In general, the organization of modules and domains corresponds exactly to the chemical structure of the initial polyketide product. This colinear property has made modular PKSs attractive subjects for rational bioengineering to produce novel bioactive compounds,[6] and has facilitated the "mining" of genome sequences for novel biosynthetic pathways.[7–9] However, there is growing interest in exceptions to this rule. An early example is the pikromycin PKS, which produces both a 12-membered and a 14-membered macrolide through optional "skipping"[10] of a PKS module. In other examples, a single extension module catalyzes two or more successive (and identical) rounds of polyketide carbon skeleton, either to give aberrant products[11] or as part of the normal biosynthetic pathway.[12–16] These exceptions promise to increase our understanding of the natural evolution of modular PKSs, and potentially offer lessons for the effective engineering of these systems.

The 36-membered macrocyclic antifungal azalomycins (AZLs) are the main products of *Streptomyces* sp. 211726, which has been isolated from mangrove rhizosphere soil.[17,18] These compounds show broad-spectrum antimicrobial activity, and cytotoxicity against a human colon tumor cell line.[17,18] Azalomycin F5a and certain F5a derivatives also show anti-MRSA activity.[19] Azalomycins F3a, F4a, and F5a are also produced by *Streptomyces malaysiensis* DSM 4137,[20] and we have previously proposed a model in which the first extension module catalyzes both the first and second cycles of polyketide chain extension, based on the observation that when the cluster is transplanted into a heterologous strain, azalomycins are produced, which rules out the participation of an additional PKS encoded elsewhere in the *S. malaysiensis* genome.[21] Whole-genome sequencing of *Streptomyces* sp. 211726 revealed a cluster (GenBank accession number: KY484834) that is highly similar, both in size and organization, to the azalomycin cluster in DSM 4137.[20] It has exactly the same module and domain arrangement and flanking auxiliary genes spanning an approximately 130 kb region of DNA (Figure 1 and Table S1 in the Supporting Information), and deletion of the entire region (Figure S1 in the Supporting Information) led to total loss of AZL production. As before, the loss of colinearity is readily localized to modules 1 and 2 (Figure 2 and Figure S2).

In order to test the hypothesis that extension-module 1 acts iteratively and supplies the missing full reduction module governing the second extension, a site-specific mutation of the essential histidine residue to alanine was introduced into the active site of the DH_1 domain in vivo (Figure S3). Liquid chromatography electrospray ionization high-resolution mass spectrometry (LC-ESI-HRMS) analysis showed that the

[*] W. Xu,[+] G. Zhai,[+] Dr. Y. Liu, Y. Li, Y. Shi, Prof. K. Hong, Prof. Z. Deng, Prof. Y. Sun
 Key Laboratory of Combinatorial Biosynthesis and Drug Discovery (Wuhan University), Ministry of Education, and
 School of Pharmaceutical Sciences, Wuhan University
 185 East Lake Road, Wuhan 430071 (P.R. China)
 E-mail: yhsun@whu.edu.cn

 Dr. H. Hong, Prof. P. F. Leadlay
 Department of Biochemistry, University of Cambridge
 80 Tennis Court Road, Cambridge CB2 1GA (UK)

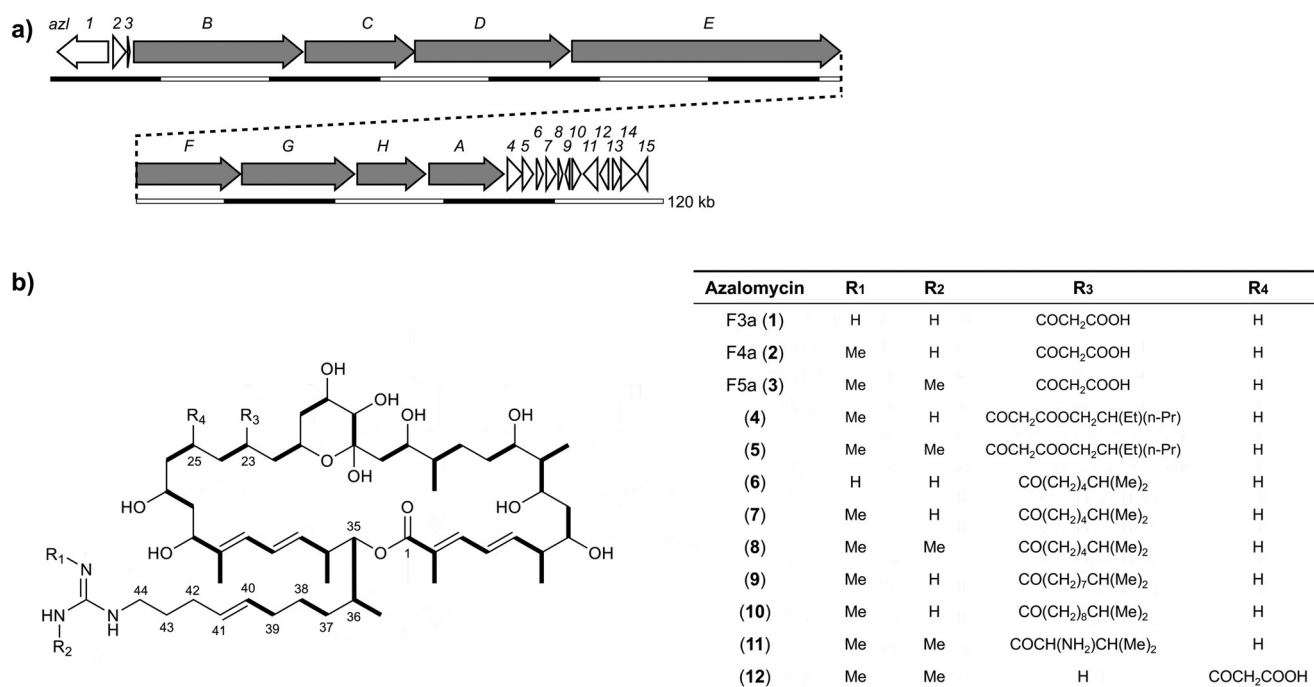[+] These authors contributed equally to this work.

Figure 1. a) Organization of the azalomycin F biosynthetic gene cluster. The PKS-encoding genes are highlighted in gray. b) Structures of azalomycin F analogues. Bold lines indicate the malonyl and methylmalonyl extender units incorporated by each acyltransferase domain of the PKS. Compounds **4–12** are minor components.

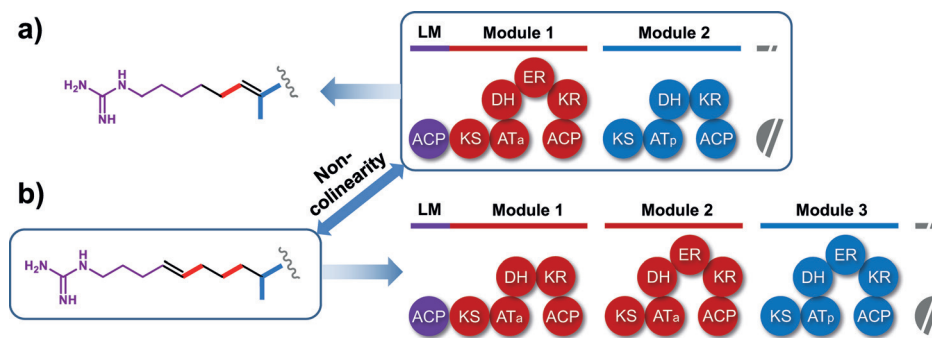| Azalomycin	R_1	R_2	R_3	R_4
F3a (**1**)	H	H	$COCH_2COOH$	H
F4a (**2**)	Me	H	$COCH_2COOH$	H
F5a (**3**)	Me	Me	$COCH_2COOH$	H
(**4**)	Me	H	$COCH_2COOCH_2CH(Et)(n\text{-}Pr)$	H
(**5**)	Me	Me	$COCH_2COOCH_2CH(Et)(n\text{-}Pr)$	H
(**6**)	H	H	$CO(CH_2)_4CH(Me)_2$	H
(**7**)	Me	H	$CO(CH_2)_4CH(Me)_2$	H
(**8**)	Me	Me	$CO(CH_2)_4CH(Me)_2$	H
(**9**)	Me	H	$CO(CH_2)_7CH(Me)_2$	H
(**10**)	Me	H	$CO(CH_2)_8CH(Me)_2$	H
(**11**)	Me	Me	$COCH(NH_2)CH(Me)_2$	H
(**12**)	Me	Me	H	$COCH_2COOH$

Figure 2. Non-colinearity between the AZL PKS assembly line and the product structure. a) Proposed chemical structure of AZL (the guanidino-substituted chain is shown) deduced from the nucleotide sequence and the predicted domain organization of the AZL biosynthetic gene cluster. b) Bioinformatic prediction of PKS organization based on the actual structure of AZL according to the canonical colinearity rule.

the DEPT analysis showed a lack of two olefinic carbon atoms at δ_C 128.9 and δ_C 131.0 Figure S6). Clearly, inactivation of the DH_1 domain prevents dehydration in both the first and second extensions, thus providing direct evidence that in the AZL PKS, module 1 is used twice.

The PKS AzlA houses only the loading module (LM) and extension-module 1. To further confirm the above result by in vitro analysis, the recombinant PKS $AzlA(\Delta DH_1)$, which bears the same site-directed mutation in the DH_1 domain as above, was successfully expressed in soluble form in *Escherichia coli* BAP1[22] with post-translational modification of the ACP_1 domain. In a one-pot reaction with 4-guanidinobutyric acid, purified Azl4 (ligase), Azl5 (acyltransferase), and $AzlA(\Delta DH_1)$, the starter unit 4-guanidino-butyryl-CoA was produced and loaded onto the ACP_L domain that comprises the LM in $AzlA(\Delta DH_1)$. In order to achieve multiple turnovers of polyketide synthesis on the AzlA PKS multienzyme, our strategy[23] was to supply additional copies of the ACP_1 domain. Recombinant *holo*-ACP_1 domain was expressed as a stand-alone protein and purified from *E. coli* BAP1 and added to the incubations to compete with the integral ACP_1 domain of module 1 (Figure S7). The polyketide-chain-bearing ACP_1 reaction products were monitored by LC-ESI-HRMS. In the presence of *holo*-ACP_1 and malonyl-CoA in the above one-pot reaction, a peak with a mass corresponding to the expected product of two rounds

ΔDH_1 mutant strain no longer produces known AZLs, and instead significant new peaks were found (Figure 3c and Figure S4), albeit at levels only 0.1% of AZL levels in the wild-type. Analysis of these new peaks revealed compounds closely related to azalomycins F3a, F4a, and F5a, except that they bear hydroxy groups at C-41 and C-39, instead of a double bond between C-40 and C-41 (Figure 3c and Figure S4). These compounds, named F3a′, F4a′, and F5a′, were isolated from a large-scale fermentation, and their structures were elucidated by ¹H NMR and DEPT spectroscopy. The NMR data showed in each case the loss of two olefinic protons (δ_H 5.38–5.46, 2H) and the presence of two additional O-bound protons at δ_H 3.65 and δ_H 4.04 (Figure S5) compared to the known azalomycins. In agreement with this,
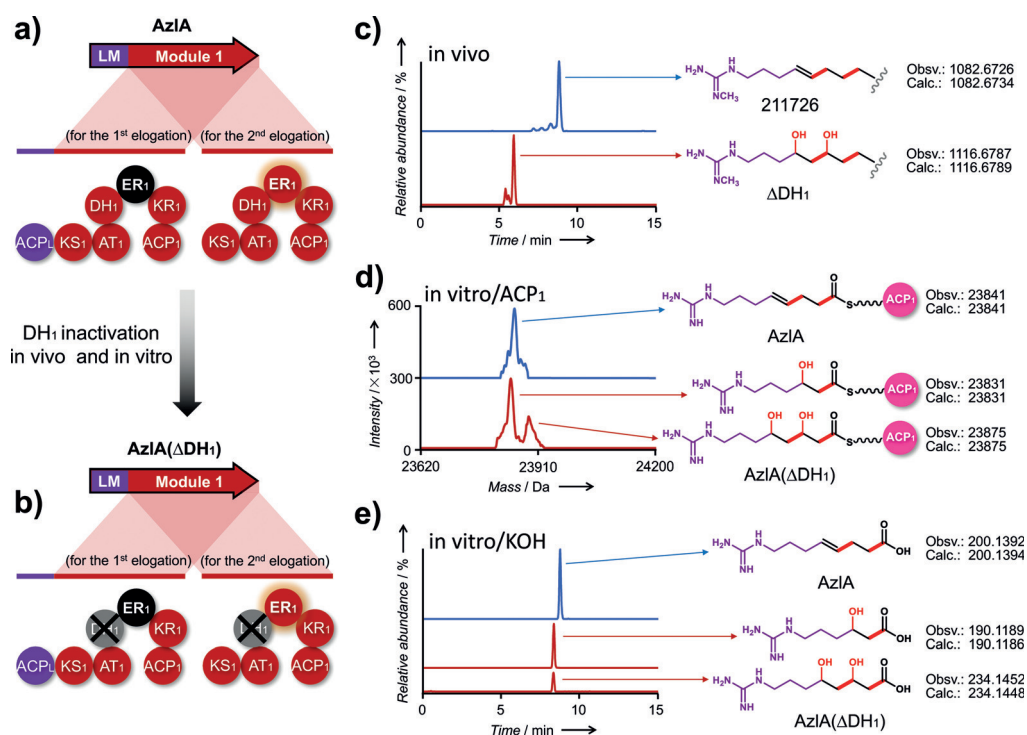
Figure 3. A model for the toggling ER₁ domain in iterative-module 1 based on the results of inactivation of the DH₁ domain in vivo and in vitro. a) Domain organization of AzlA. Module 1 is used twice, and the toggling ER₁ domain is programmed to be inactive (shaded in black) in the first extension but active in the second extension. b) Domain organization of AzlA(ΔDH₁). The DH₁ domain marked with a cross was inactivated by site-specific mutagenesis in the 211726 chromosome and the AzlA recombinant multienzyme. c) LC-ESI-HRMS analysis of fermentation products from wild-type 211726 and the ΔDH₁ mutant. Only F4a and its derivative F4b, the main products of AZL, are shown. d) LC-ESI-HRMS protein analysis of an in vitro reconstitution assay of AzlA and AzlA(ΔDH₁) with added recombinant stand-alone ACP₁. e) LC-ESI-HRMS analysis of the same in vitro assay mixture as in (d) except that the reaction mixture was subjected to alkaline hydrolysis before analysis. The three traces are presented on the same scale.

a one-pot reaction with wild-type AzlA (Figure S7), was used. Since module 1 contains a full set of reductive enzyme domains (DH, ER, KR), and recruits malonyl-CoA as extender unit, the 4-guanidinobutyryl starter unit should be extended with two acetate units and both newly-formed β-keto groups should be fully reduced. However, LC-ESI-HRMS analysis of the one-pot reaction showed that the molecular weight of the acyl-ACP₁ species produced was 2 Da less than expected for the product of two rounds with full reduction (Figure 3 d), thus suggesting incomplete reduction in one cycle. To confirm this, analysis of the acyl chain after its release from ACP₁ by alkaline hydrolysis showed it to be 8-guanidinooct-4-enoic acid (Figure 3 e). Therefore, AzlA catalyzes two elongations, in only one of which the enoylreductase (ER) domain acts. The position of the double bond in all natural AZL compounds (between C-40 and C-41) is fully consistent with ER₁ being inactive in the first extension but active in the second one.

of extension (8-guanidino-3,5-dihydroxyoctanoyl-ACP₁) was observed (Figure 3 d). A further peak was detected corresponding to the mass of 6-guanidino-3-hydroxyhexanoyl-ACP₁, the expected product of a single cycle of chain extension (Figure 3 d). To confirm these findings, thioester-bound products were hydrolyzed with potassium hydroxide as described[24] in both assays with and without *holo*-ACP₁ and showed the same result. LC-ESI-HRMS analysis of the hydrolyzed products revealed two peaks, corresponding to 6-guanidino-3-hydroxyhexanoic acid and 8-guanidino-3,5-dihydroxyoctanoic acid, respectively, which is in agreement with the analysis of the acyl-ACP₁ species (Figure 3 e). Taken together, these results demonstrate conclusively that the AZL PKS extension-module 1 catalyzes two successive cycles of chain extension. In the iterative mechanism we propose for the intact AZL PKS (Figure S7), after the first round of extension, the ACP₁-tethered polyketide intermediate is transferred back to the KS₁ domain on the opposite PKS strand[25] instead of the downstream KS₂ domain, probably owing to the preference of KS₂ for the triketide acyl-ACP as a substrate over the diketide acyl-ACP.

Notably, only one full reduction occurs during the first two extension cycles on the AZL PKS. To investigate the behavior of the ER₁ domain in module 1, the same experimental approach as above, in which recombinant ACP₁ was used in

only one of which the enoylreductase (ER) domain acts. The position of the double bond in all natural AZL compounds (between C-40 and C-41) is fully consistent with ER₁ being inactive in the first extension but active in the second one.

The programmed iteration of extension-module 1 on the AZL PKS occurs, as with previously identified examples of programmed iteration of modules,[11,26] at a point in the assembly line where there is a protein–protein interface[27] (here between AzlA and AzlB). Those previous examples all involve the same level of β-keto processing in each of the iterative rounds of chain elongation. Here, we report an unprecedented example in which every domain except the ER is used during two successive rounds of elongation, while the ER domain is non-functional ("off") in the first but functional ("on") in the second extension process. Nevertheless, the iterative use of module 1 and the ER₁ skipping event in the first extension can both be accounted for in terms of kinetic control: the intrinsic selectivity of the ER₁ domain, which discriminates against the shorter substrate, would favor back-transfer of the acyl chain onto KS₁. Likewise, for the KS₁ active site to be re-used, the back-transfer must outcompete the priming of KS₁ through transfer of the 4-guanidinobutyryl starter unit from the adjacent ACP₁ loading domain. After two rounds of extension, the partitioning of the triketide acyl-ACP intermediate favors full reduction rather than transfer to

the next module. The downstream KS$_2$ domain on the adjacent PKS subunit AzlB may well act as a "gatekeeper", favoring selective recruitment of a triketide- rather than a diketide-acyl chain. Evidence from a study of the aureothin PKS[28] strongly supports such an interplay of multiple factors in programmed iteration.

Selective processing during two successive iterations of polyketide chain extension is a known feature in highly reducing fully iterative polyketide synthases.[29–31] In the LovB PKS, for example, which synthesizes the nonaketide core of the cholesterol-lowering compound lovastatin using a single PKS module, an integral methyltransferase domain selectively methylates the β-ketothioester formed after three rounds of chain extension.[29] In vitro analysis of model thioester substrates for LovB has demonstrated that the methyltransferase is exquisitely specific for the tetraketide substrate, and that for this substrate, methylation effectively outcompetes prior ketoreduction.[32] LovB also recruits an exogenous monomeric ER protein (LovC) to catalyze specific enoyl reduction in only three out of the eight extension cycles (tetra-, penta-, and heptaketide intermediates).[32] Detailed analysis of LovC has suggested that while the competent substrates are readily accommodated in a productive conformation, shorter intermediates preferentially adopt nonproductive conformations in the active site.[32] Meanwhile, the hexaketide intermediate preferentially undergoes Diels–Alder-like cyclization. Similar experiments will be needed to establish the precise structural basis for the remarkable switch that flips the AzlA ER$_1$ domain from "off" to "on" during azalomycin biosynthesis.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

[1] J. Cortés, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **1990**, *348*, 176–178.

[2] S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **1991**, *252*, 675–679.

[3] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380–416.

[4] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468–3496.

[5] C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716; *Angew. Chem.* **2009**, *121*, 4782–4811.

[6] K. J. Weissman, P. F. Leadlay, *Nat. Rev. Microbiol* **2005**, *3*, 925–936.

[7] A. Starcevic, J. Zucko, J. Simunkovic, P. F. Long, J. Cullum, D. Hranueli, *Nucleic Acids Res.* **2008**, *36*, 6882–6892.

[8] M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, *Nucleic Acids Res.* **2011**, *39*, W339–46.

[9] M. H. Li, P. M. Ung, J. Zajkowski, S. Garneau-Tsodikova, D. H. Sherman, *BMC Bioinf.* **2009**, *10*, 185.

[10] B. J. Beck, Y. J. Yoon, K. A. Reynolds, D. H. Sherman, *Chem. Biol.* **2002**, *9*, 575–583.

[11] B. Wilkinson, G. Foster, B. A. Rudd, N. L. Taylor, A. P. Blackaby, P. J. Sidebottom, D. J. Cooper, M. J. Dawson, A. D. Buss, S. Gaisser, et al., *Chem. Biol.* **2000**, *7*, 111–117.

[12] N. Gaitatzis, B. Silakowski, B. Kunze, G. Nordsiek, H. Blöcker, G. Höfle, R. Müller, *J. Biol. Chem.* **2002**, *277*, 13082–13090.

[13] J. He, C. Hertweck, *Chem. Biol.* **2003**, *10*, 1225–1232.

[14] J. He, C. Hertweck, *ChemBioChem* **2005**, *6*, 908–912.

[15] C. Olano, B. Wilkinson, S. J. Moss, A. F. Braña, C. Méndez, P. F. Leadlay, J. A. Salas, *Chem. Commun.* **2003**, 2780–2782.

[16] C. Olano, B. Wilkinson, C. Sánchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Braña, C. J. Martin, M. Oliynyk, et al., *Chem. Biol.* **2004**, *11*, 87–97.

[17] G. Yuan, H. Lin, C. Wang, K. Hong, Y. Liu, J. Li, *Magn. Reson. Chem.* **2011**, *49*, 30–37.

[18] G. Yuan, K. Hong, H. Lin, Z. She, J. Li, *Mar. Drugs* **2013**, *11*, 817–829.

[19] G.-J. Yuan, P.-B. Li, J. Yang, H.-Z. Pang, Y. Pei, *Chin. J. Nat. Med.* **2014**, *12*, 309–313.

[20] H. Hong, T. Fill, P. F. Leadlay, *Angew. Chem. Int. Ed.* **2013**, *52*, 13096–13099; *Angew. Chem.* **2013**, *125*, 13334–13337.

[21] H. Hong, Y. Sun, Y. Zhou, E. Stephens, M. Samborskyy, P. F. Leadlay, *Beilstein J. Org. Chem.* **2016**, *12*, 2164–2172.

[22] B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **2001**, *291*, 1790–1792.

[23] Y. Sun, F. Hahn, Y. Demydchuk, J. Chettle, M. Tosin, H. Osada, P. F. Leadlay, *Nat. Chem. Biol.* **2009**, *6*, 99–101.

[24] S. E. O'Connor, H. Chen, C. T. Walsh, *Biochemistry* **2002**, *41*, 5685–5694.

[25] B. Busch, N. Ueberschaar, Y. Sugimoto, C. Hertweck, *J. Am. Chem. Soc.* **2012**, *134*, 12382–12385.

[26] H. Chen, L. Du, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 541–557.

[27] S. J. Moss, C. J. Martin, B. Wilkinson, *Nat. Prod. Rep.* **2004**, *21*, 575–519.

[28] B. Busch, N. Ueberschaar, S. Behnken, Y. Sugimoto, M. Werneburg, N. Traitcheva, J. He, C. Hertweck, *Angew. Chem. Int. Ed.* **2013**, *52*, 5285–5289; *Angew. Chem.* **2013**, *125*, 5393–5397.

[29] J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. R. Hutchinson, *Science* **1999**, *284*, 1368–1372.

[30] L. M. Halo, J. W. Marshall, A. A. Yakasai, Z. Song, C. P. Butts, M. P. Crump, M. Heneghan, A. M. Bailey, T. J. Simpson, C. M. Lazarus, et al., *ChemBioChem* **2008**, *9*, 585–594.

[31] K. Belecki, J. M. Crawford, C. A. Townsend, *J. Am. Chem. Soc.* **2009**, *131*, 12564–12566.

[32] B. D. Ames, C. Nguyen, J. Bruegger, P. Smith, W. Xu, S. Ma, E. Wong, S. Wong, X. Xie, J. W. H. Li, et al., *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11144–11149.