Polyketide Biosynthesis

DOI: 10.1002/anie.201406602

# An Unusual Dehydratase Acting on Glycerate and a Ketoreducatse Stereoselectively Reducing α-Ketone in Polyketide Starter Unit Biosynthesis**

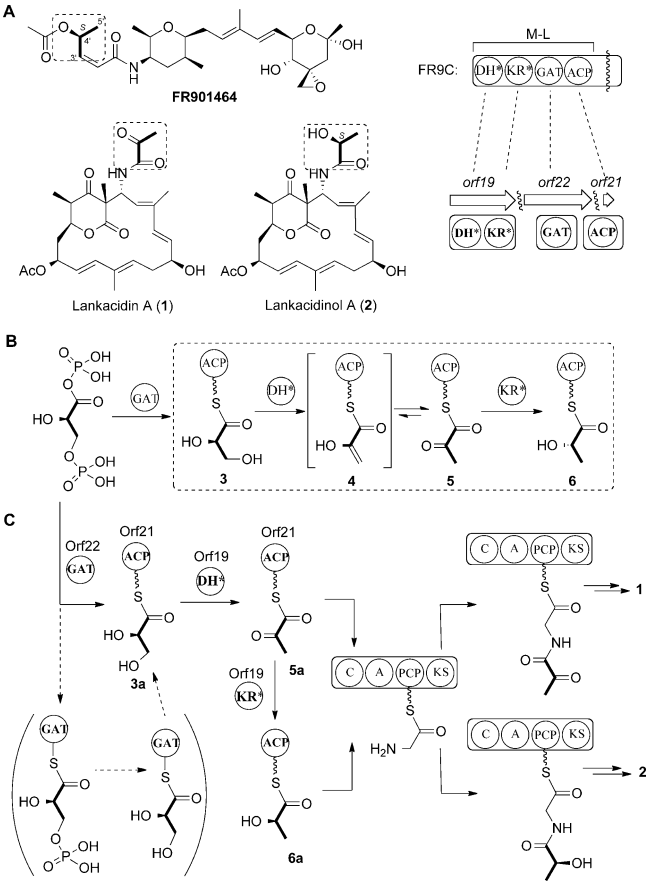
*Hai-Yan He, Hua Yuan, Man-Cheng Tang, and Gong-Li Tang**

*Dedicated to Professors Chengye Yuan and Lixin Dai on the occasion of their 90th birthdays*

***Abstract:*** *Polyketide synthases (PKSs) usually employ a ketoreductase (KR) to catalyze the reduction of a β-keto group, followed by a dehydratase (DH) that drives the dehydration to form a double bond between the α- and β-carbon atoms. Herein, a DH*-KR* involved in FR901464 biosynthesis was characterized: DH* acts on glyceryl-S-acyl carrier protein (ACP) to yield ACP-linked pyruvate; subsequently KR* reduces α-ketone that yields* L*-lactyl-S-ACP as starter unit for polyketide biosynthesis. Genetic and biochemical evidence was found to support a similar pathway that is involved in the biosynthesis of lankacidins. These results not only identified new PKS domains acting on different substrates, but also provided additional options for engineering the PKS starter pathway or biocatalysis.*

Modular polyketide synthases (PKSs), also known as type I PKSs, have been well-established to catalyze the biosynthesis of a large group of polyketide natural agents with remarkably structural diversity using a thiotemplated assembly line. Each module contains a β-ketoacyl synthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP), which catalyze one cycle of chain elongation that extend the growing polyketide chain by a $C_2$ unit.[1] Other chemically diverse polyketide skeletons arise from processing domains, such as ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains, which may be present in a module in various combinations to control the oxidation state and stereochemistry of the growing polyketide chain.[1] The DH and KR domains, including their structure-based catalytic mechanism,[2] regio- and stereoselectivity,[3] and biocatalytic potency,[4] have been well-studied. Most KRs utilize NADPH to stereoselectively reduce the β-keto group of a β-ketoacyl-S-ACP intermediate generated by the KSs; then, DHs catalyze the dehydration of the resulting intermediate to form a double bond between the α- and β-carbons.

FR901464 (Scheme 1 A) is a natural antitumor agent and representing a new class of potent anticancer small molecules that target the spliceosome inhibiting both splicing and nuclear retention of pre-mRNA.[5] Previous studies have




***Scheme 1.*** Loading modules of PKSs in FR901464 and lankacidin biosynthesis. A) The molecular structure and the loading module (M-L). DH, dehydratase; KR, ketoreductase; GAT (or FkbH), glyceryl transferase/phosphatase; ACP, acyl carrier protein. The three-carbon starter units are boxed. B) Proposed reactions catalyzed by the M-L, the reactions by DH*-KR* are highlighted in a box. C) Proposed starter unit biosynthesis of lankacidin (**1**) and lankacyclinol (**2**).

[*] Dr. H.-Y. He, Dr. H. Yuan, Dr. M.-C. Tang, Prof. Dr. G.-L. Tang
State Key Laboratory of Bioorganic and Natural Products Chemistry
Shanghai Institute of Organic Chemistry
Chinese Academy of Sciences
345 Lingling Road, Shanghai, 200032 (China)
E-mail: gltang@sioc.ac.cn
Prof. Dr. G.-L. Tang
Shanghai Collaborative Innovation Center for Biomanufacturing Technology
130 Meilong Road, Shanghai 200237 (China)

[**] We thank Prof. Zixin Deng's Lab at Shanghai JiaoTong University for support in obtaining MS data. This work was financially supported by grants from the 973 Program (2010CB833200) and NNSFC (31200054 and 31330003).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201406602.

revealed that it is biosynthesized by "*trans*-AT" PKSs hybridized with nonribosomal peptide synthetase (NRPS) and an isoprenoid-like β-branching pathway.[6] In the loading module, a DH- and a KR-like didomain (DH*-KR*, that did not have an exact match throughout the sequence) were proposed to catalyze the biosynthesis of unusual PKS starter unit (Scheme 1 A).[6] Although a similar domain organization and suggested biosynthetic pathway were also described in the biosynthesis of bryostatin (Supporting Information, Figure S1), an anticancer polyketide from an uncultivated bacterial symbiont, based on the bioinformatic analysis,[7] the enzymatic logic has remained unknown. Herein, we report the biochemical elucidation of the physiological role of this type of novel domains: DH* first acting on glycerate; KR* then stereoselectively reducing α-ketone, which differs from known PKSs.

It has been established that a glyceryl transferase (GAT) and an ACP in the loading module incorporate D-1,3-bisphosphoglycerate (1,3-BPG) to afford ACP-tethered glycerate **3**.[6] The following enzymatic pathway has been proposed: first, DH* catalyzes the dehydration of **3** to an ACP-linked enoylpyruvate **4**, which could spontaneously rearrange to form ACP-bound pyruvate **5**; then, KR* carries out α-ketone reduction to yield L-lactyl-*S*-ACP **6**, which serves as the starter unit for PKS in FR901464 biosynthesis (Scheme 1 B).[7] To validate this postulation, the DH* domain was firstly expressed and purified from *E. coli* BL21 (DE3) (Supporting Information, Figure S2). Considering that the retention time of **3** and holo-ACP is almost the same in high-performance liquid chromatography (HPLC) analysis,[6] we chose another ACP-tethered glycerate **3b** (Figure 1 A, with a different retention time to the respective holo-ACP), from the quartromicin (QMN) biosynthetic pathway,[8] as a substrate mimic to facilitate detection. After **3b** was generated in situ catalyzed by QmnD1 (a GAT coupled with QmnD2 in QMN biosynthesis, Figure 1 B-I), it was incubated with Fr9C-DH* and the reaction products were then subjected to HPLC analysis. The result shows that the peak height of **3b** decreased and that of holo-ACP increased, but no new peak signals were detected (Figure 1 B-II). In a control assay containing the boiled enzyme, the substrate **3b** did not show any changes over several hours (Figure 1 B-III). When the reaction products were detected using high-resolution MS (HRMS), a trace of a new protein possessing identical molecule weight with ACP-bound pyruvate **5b** was observed (Figure 1 C-III). The α-keto ester **5b** is difficult to detect owing to its instability towards hydrolysis to give the α-keto acid **5c** and holo-ACP (Figure 2 A), as described by Calderone.[9] To solve this problem, we added derivative reagent 4-dinitrophenylhydrazine (DNPH) to the reactions to detect α-keto acid (Figure 2 A).[10] As expected, pyruvate derivative **5d** was observed in the Fr9C-DH* assay only (Figure 2 B). These results indicate that Fr9C-DH* catalyzes β-dehydration of ACP-tethered glycerate **3b** to yield pyruvoyl-*S*-ACP **5b** (Figure 1 A).

Next, we investigated the function of Fr9C-KR*. DH*-KR* didomain protein was expressed and purified for biochemical assays, since expression of the independent KR* domain was not successful owing to its insolubility in

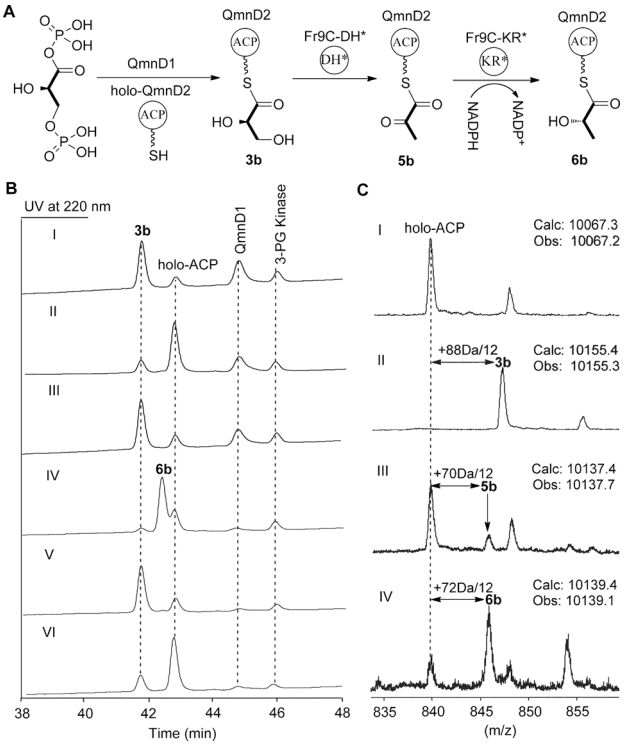



***Figure 1.*** Biochemical characterization of Fr9C-DH*-KR*. A) Reactions in these assays. B) HPLC analysis: I, generation of **3b** in situ; II, **3b** with Fr9C-DH*; III, **3b** with boiled Fr9C-DH*; IV, **3b** with Fr9C-DH*-KR* and NADPH; V, **3b** with Fr9C-DH*-KR* and NADH for 30 min; VI, **3b** with Fr9C-DH*-KR* and NADH for 2 h. C) HRMS analysis: I, holo-ACP; II, generation of **3b**; III, **3b** with Fr9C-DH*; IV, **3b** with Fr9C-DH*-KR* and NADPH.

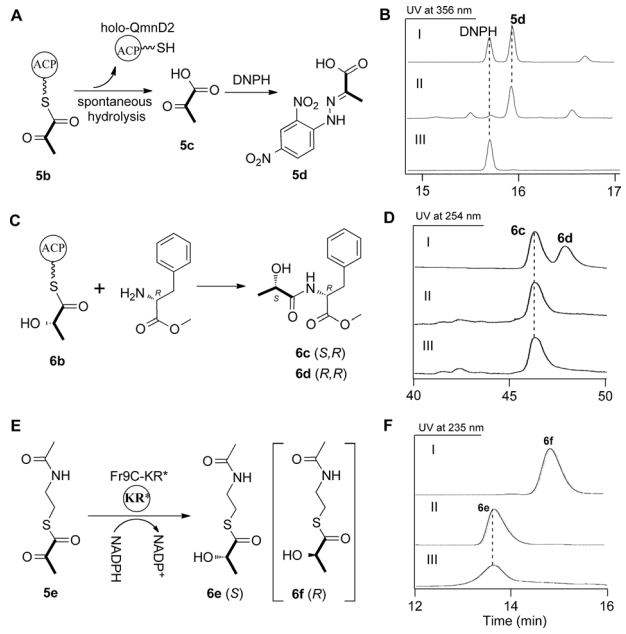



***Figure 2.*** Chemical derivatives and assays of KR*. A) The derivative reaction of pyruvate. B) HPLC analysis: I, standard control; II, full assay; III, assay with boiled Fr9C-DH*. C) The derivative reaction of lactyl unit. D) HPLC analysis: I, standard control; II, assay with Fr9C-DH*-KR* and NADPH; III, assay with Orf19 and NAD(P)H. E) Reaction catalyzed by KR. F) HPLC analysis: I and II, standard; III, full assay with Fr9C-DH*-KR* and NADPH.

*E. coli* BL21 (DE3) (Supporting Information, Figure S2). Similarly, after **3b** was generated by in situ, DH*-KR* and cofactor NADPH were added to the reaction. HPLC analysis revealed that the amount of **3b** decreased and a new peak emerged (Figure 1 B-V and VI). HRMS analysis reported that the molecular weight of this new peak is consistent with lactyl-*S*-ACP **6b** (Figure 1 C-IV); however, when NADPH was replaced by NADH, production of **6b** was not observed, even though the reaction lasted 2 h. Most of **3b** had been converted, which is similar to the result of only adding Fr9C-DH* (Figure 1 B-IV). Furthermore, if D-phenylalanine methyl ester as derivative reagent was added to the Fr9C-DH*-KR* assays, it will react with **6b** to give **6c** (Figure 2 C).[11] We chemically synthesized diastereoisomer **6c** and **6d** for potential product standards (Supporting Information, Figure S3). As expected, generation of **6c**, the lactyl chirality of which is consistent with that of final product FR901464, was detected in the derivative assay, but **6d** with opposite chirality cannot be observed (Figure 2 D-II). We also synthesized *N*-acetyl cysteamine (NAC) thioester **5e** to mimic the ACP-bound substrate **5b** of KR* domain and stereoisomer **6e** and **6f** for potential product standard (Figure 2 E). After incubation of **5e** with NADPH and the DH*-KR*, HPLC and MS analyses showed that the anticipated reductive product **6e** could be detected, but **6f** with opposite chirality was not observed (Figure 2 F); though, owing to instability of **5e** in its aqueous phase, conversion efficiency of **5e** to **6e** was very low (about 5–10 %). These results demonstrated that the Fr9C-KR* domain could use NADPH, but not NADH, as a hydrogen donor to reduce the pyruvoyl unit to L-lactyl unit tethered on ACP stereospecifically (Figure 1 A). Furthermore, consumption of the unstable intermediate pyruvoyl-*S*-ACP **5b** into **6b** by KR* could accelerate the dehydration by DH*, which hints at the functional correlation between two domains.

DH*-KR*-GAT-ACP organization is rarely used as a PKS starter module for the biosynthesis of polyketide natural products. Other than bryostatin, additional such cases can be found in thailanstatin,[12] the symmetric polyketide dimer SIA7248[13] and the tartrolons[14] biosynthetic pathways (Supporting Information, Figure S1). Additionally, a homologous sequence was discovered in Genbank, which is located within a linear plasmid in *Streptomyces rochei* 7434AN4.[15] This homologous sequence contains three independent genes, *orf19*, *orf21* and *orf22*, which encode a DH*-KR* didomain protein (Orf19), an ACP (Orf21) and a GAT-like protein (Orf22), respectively (Scheme 1 A). These genes are adjacent to the gene cluster of lankacidin which is biosynthesized by a hybrid PKS/NRPS system.[16] The structural difference between lankacidin A (**1**) and its analogue lankacyclinol A (**2**) is that the three-carbon unit is linked to the amino group with pyruvoyl and L-lactyl group, respectively (Scheme 1 A). Although two groups have reported the biosynthetic studies of **1** and **2**,[16] the origin and pathway of the three-carbon unit is still unknown. Considering the FR901464 starter pathway, we believed that the genes *orf19*, *orf21*, and *orf22* may be related to the biosynthesis of the three-carbon units of **1** and **2** (Scheme 1 C). This proposal had also been discussed by Sherman and Haygood in reporting on the bryostatin gene cluster;[7] however, there is no any genetic or biochemical evidence to support this hypothesis to date. Thus, in vivo experiments were carried out and we found that production of **1** and **2** were terminated completely in knockout mutants of *orf19* and *orf22*. When *orf19* and *orf22* were complemented to the mutant respectively, production of **1** and **2** was regained, though the yield was lower than that of the wild type (Supporting Information, Figure S4 and S5). These results indicate that *orf19* and *orf22* are essential for the biosynthesis of **1** and **2**. Next, we used in vitro analyses to verify our speculation. Three proteins Orf19, Orf21, and Orf22 were expressed and purified in *E. coli* BL21 (DE3) (Supporting Information, Figure S2). Similily, apo-Orf21 was converted completely into active holo-Orf21 by Sfp (Figure 3 A-I and 3 B-I). Then 1,3-BPG and bifunctional glyceryl transferase/

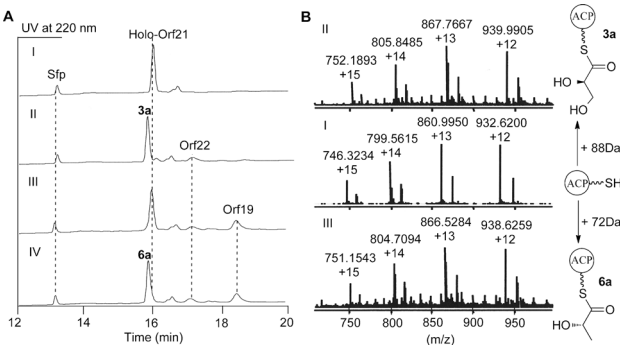



***Figure 3.*** Biochemical characterization of Orf19, Orf21, and Orf22. A) HPLC analysis: I, generation of holo-Orf21; II, generation of **3a** catalyzed by Orf 22; III, assay of **3a** with Orf19; IV, full assay of **3a** with Orf19 and NAD(P)H. B) HRMS analysis: I, holo-ACP; II, generation of **3a**; III, assay of **3a** with Orf19 and NAD(P)H.

phosphatase Orf22 were added to the reaction. Glyceroyl-*S*-ACP **3a** was detected (Scheme 1 C, Figure 3 A-II and 3 B-II), illustrating that Orf22 is fully functional. The DH*-KR* didomain protein Orf19 was next added into the reaction using **3a** as substrate. As expected, the peak of **3a** disappeared and the amount of holo-Orf21 increased (Figure 3 A-III), which is analogous to the phenomena observed in the dehydration reaction. When NADPH or NADH were added as hydrogen donors, a new peak emerged with molecular weight consistent with that of **6a** (Figure 3 A-IV and B-III); this assay was also dealt with derivative reagent, product **6c** was detected (Figure 2 D-III). By combining in vivo and in vitro results, we established the essential roles of *orf19*, *orf21*, and *orf22* in the biosynthesis of lankacidins and identified the missing link of the starter pathway of **1** and **2** (Scheme 1 C).

In modular PKS systems, DH and KR domains are essential components for the structural diversity of various products.[1] Recent structural studies have revealed that Asp residue donates a proton to the β-hydroxy group and His residue abstracts an α-proton in the DH domain, which play vital roles in dehydration, resulting in an α,β-unsaturated intermediate.[2] DH domains have two hotdog folds and the His residue is located on the small cap of the N-terminal; the

Asp is on a helix of the C-terminal hotdog.[17] In the present work, it was confirmed that the two DH*-like proetins (Fr9C-DH* and Orf19-DH*) function as DHs acting on the glyceryl-*S*-ACP substrate; however, further bioinformatic analysis of these DH*-like enzymes suggested that these proteins belong to a family of MaoC-like dehydratases (pfam 01575, Supporting Information, Figure S6), which usually act as (*R*)-specific enoyl-CoA hydratase and have a conserved catalytic DxxxxH motif.[18] The Asp residue activates a water to attack the C3 carbon of the substrate, and the His residue donates a proton to the C2 carbon. Interestingly, multiple sequence alignment analysis revealed that the DH* sequences also have a conserved DxxxxH motif (Supporting Information, Figure S6). Considering the reversible reaction properties of these enzymes,[19] we inferred that this motif is also essential for the dehydration activity in Fr9C-DH* and Orf19-DH*. As expected, biochemical assays showed that both the His and Asp mutants lost their dehydration activity (Supporting Information, Figure S8). Thus, the DH*-like domains may be a new family of DHs that could be integrated into the PKS modules and have the catalytic DxxxxH motif as in (*R*)-specific enoyl-CoA hydratases but catalyze the β-dehydration of glyceryl-*S*-ACP (Figure 4A).

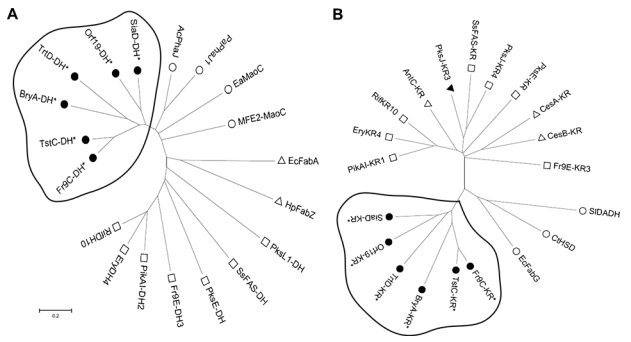



***Figure 4.*** A neighbor-joining tree cladogram of DHs (A) and KRs (B). A) DH*-like (●); MaoC-family (○); FabA-family (△); DHs involved in PKSs and fatty acid synthases (□). B) KR*-like (●); SDR-family (○); KRs involved in PKSs and fatty acid synthases (□); α-KRs involved in NRPSs (△); α/β-KR in PksJ (▼). The details for every gene/protein are provided in the Supporting Information.

Unlike β-KRs, α-KRs are rare in PKS systems. In the biosynthesis of bacillaene, a KR domain in PksJ is thought to be a bifunctional enzyme that could catalyze both α- and β-ketoreduction.[9] Other reports have revealed that such α-KRs are either integrated into NRPS loading modules with domain organization of A-KR-PCP (cereulide and valinomycin)[20] or integrated into an extension module with C-A-KR-PCP (antimycin).[21] Sequence analysis showed that these α-KRs and PKS β-KRs belong to the same family of the KR domain (PF08659), while the present KR* belong to the family of short-chain dehydrogenase/reductases (SDR, PF00106). Previous studies have established that the SDRs possess a catalytic tetrad of N-S-Y-K residues,[22] and multiple alignment revealed these KR*s have a similar conserved S-Y-K catalytic triad (Supporting Information, Figure S7). As expected, mutation of the key Tyr430 into Ala abolished the activities of Fr9C-KR* (Supporting Information, Figure S8). Further neighbour-joining tree analysis revealed that these KR*s could also be classified into a new family of KRs, although they have sequence homology to SDR (Figure 4B). It is well-established that β-KRs can be classified into A- and B-types according to their sequences, which catalyze the opposite configuration of products.[23] Owing to the structure of the lactyl unit, we inferred that α-KR* that play a role in the biosynthesis of FR901464, thailanstatins, lankacinins, SIA7248, and tartrolons are all responsible for L-OH formation, while only α-KR* that plays a role bryostatin biosynthesis produces D-OH (Supporting Information, Figure S1). However, multiple alignment analysis of these α-KR*s did not reveal any clues to predict the product configuration, which do not resemble the well studied PKS β-KRs. Several mutants of Fr9C-KR* were constructed to alter the configuration of reduction products, but all tries were unsuccessful (data not shown).

Glyceroyl-*S*-ACP is an unusual extender unit used in PKS which could be further transformed into hydroxymalonyl-*S*-ACP or methoxymalonyl-*S*-ACP.[24] Herein we discovered a new conversion by a DH*-KR* bifunctional protein into lactyl-*S*-ACP, which serves as a starter unit for PKS in FR901464 and for hybrid NRPS/PKS in lankacidin biosynthesis. Further bioinformatic analysis allowed us to identify at least seven other biosynthetic systems containing the DH*-KR* like enzymes based on the genome sequence, although the respective biosynthetic pathways or products have never been explored (Supporting Information, Table S4). Thus, our characterizations of a different starter pathway in polyketide biosynthesis not only provide a new strategy for combinatorial biosynthesis, but also could impact in the discovery of novel natural products by genome mining.

Received: June 26, 2014
Published online: August 27, 2014

**Keywords:** dehydratase · FR901464 · ketoreductase · PKS starter · α-ketone reduction

[1] a) E. S. Sattely, M. A. Fischbach, C. T. Walsh, *Nat. Prod. Rep.* **2008**, *25*, 757–793; b) C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716; *Angew. Chem.* **2009**, *121*, 4782–4811.
[2] a) A. T. Keatinge-Clay, *Nat. Prod. Rep.* **2012**, *29*, 1050–1073; b) D. Gay, Y.-O. You, A. Keatinge-Clay, D. E. Cane, *Biochemistry* **2013**, *52*, 8916–8928.
[3] a) D. E. Cane, *J. Biol. Chem.* **2010**, *285*, 27517–27523; b) D. H. Kwan, F. Schulz, *Mol.* **2011**, *16*, 6092–6115.
[4] a) S. Bali, H. M. O'Hare, K. J. Weissman, *ChemBioChem* **2006**, *7*, 478–484; b) S. Bali, K. J. Weissman, *ChemBioChem* **2006**, *7*, 1935–1942.
[5] a) H. Nakajima, B. Sato, T. Fujita, S. Takase, H. Terano, M. Okuhara, *J. Antibiot.* **1996**, *49*, 1196–1203; b) D. Kaida, H. Motoyoshi, E. Tashiro, T. Nojima, M. Hagiwara, K. Ishigami, H. Watanabe, T. Kitahara, T. Yoshida, H. Nakajima, T. Tani, S. Horinouchi, M. Yoshida, *Nat. Chem. Biol.* **2007**, *3*, 576–583.
[6] F. Zhang, H.-Y. He, M.-C. Tang, Y.-M. Tang, Q. Zhou, G.-L. Tang, *J. Am. Chem. Soc.* **2011**, *133*, 2452–2462.

[7] M. Hildebrand, L. E. Waggoner, H. Liu, S. Sudek, S. Allen, C. Anderson, D. H. Sherman, M. Haygood, *Chem. Biol.* **2004**, *11*, 1543 – 1552.
[8] H.-Y. He, H.-X. Pan, L.-F. Wu, B.-B. Zhang, H.-B. Chai, W. Liu, G.-L. Tang, *Chem. Biol.* **2012**, *19*, 1313 – 1323.
[9] C. T. Calderone, S. B. Bumpus, N. L. Kelleher, C. T. Walsh, N. A. Magarvey, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12809 – 12814.
[10] D. H. Scharf, P. Chankhamjon, K. Scherlach, T. Heinekamp, M. Roth, A. A. Brakhage, C. Hertweck, *Angew. Chem. Int. Ed.* **2012**, *51*, 10064 – 10068; *Angew. Chem.* **2012**, *124*, 10211 – 10215.
[11] F. Kopp, U. Linne, M. Oberthür, M. A. Marahiel, *J. Am. Chem. Soc.* **2008**, *130*, 2656 – 2666.
[12] X. Liu, S. Biswas, M. G. Berg, C. M. Antapli, F. Xie, Q. Wang, M.-C. Tang, G.-L. Tang, L. Zhang, G. Dreyfuss, Y.-Q. Cheng, *J. Nat. Prod.* **2013**, *76*, 685 – 693.
[13] Y. Zou, H. Yin, D. Kong, Z. Deng, S. Lin, *ChemBioChem* **2013**, *14*, 679 – 683.
[14] S. I. Elshahawi, A. E. Trindade-Silva, A. Hanora, A. W. Han, M. S. Flores, V. Vizzoni, C. G. Schrago, C. A. Soares, G. P. Concepcion, D. L. Distel, E. W. Schmidt, M. G. Haygood, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E295 – E304.
[15] S. Mochizuki, K. Hiratsu, M. Suwa, T. Ishii, F. Sugino, K. Yamada, H. Kinashi, *Mol. Microbiol.* **2003**, *48*, 1501 – 1510.
[16] a) K. Arakawa, F. Sugino, K. Kodama, T. Ishii, H. Kinashi, *Chem. Biol.* **2005**, *12*, 249 – 256; b) J. S. Dickschat, O. Vergnolle, H. Hong, S. Garner, S. R. Bidgood, H. C. Dooley, Z. Deng, P. F. Leadlay, Y. Sun, *ChemBioChem* **2011**, *12*, 2408 – 2412.
[17] J. W. Labonte, C. A. Townsend, *Chem. Rev.* **2013**, *113*, 2182 – 2204.
[18] T. Hisano, T. Tsuge, T. Fukui, T. Iwata, K. Miki, Y. Doi, *J. Biol. Chem.* **2003**, *278*, 617 – 624.
[19] Y. M. Qin, A. M. Haapalainen, S. H. Kilpeläinen, M. S. Marttila, M. K. Koski, T. Glumoff, D. K. Novikov, J. K. Hiltunen, *J. Biol. Chem.* **2000**, *275*, 4965 – 4972.
[20] N. A. Magarvey, M. Ehling-Schulz, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 10698 – 10699.
[21] a) Y. Yan, L. Zhang, T. Ito, X. Qu, Y. Asakawa, T. Awakawa, I. Abe, W. Liu, *Org. Lett.* **2012**, *14*, 4142 – 4145; b) M. Sandy, Z. Rui, J. Gallagher, W. Zhang, *ACS Chem. Biol.* **2012**, *7*, 1956 – 1961.
[22] U. Oppermann, C. Filling, M. Hult, N. Shafqat, X. Wu, M. Lindh, J. Shafqat, E. Nordling, Y. Kallberg, B. Persson, H. Jörnvall, *Chem.-Biol. Interact.* **2003**, *143 – 144*, 247 – 253.
[23] a) P. Caffrey, *ChemBioChem* **2003**, *4*, 654 – 657; b) A. T. Keatinge-Clay, *Chem. Biol.* **2007**, *14*, 898 – 908.
[24] Y. A. Chan, A. M. Podevels, B. M. Kevany, M. G. Thomas, *Nat. Prod. Rep.* **2009**, *26*, 90 – 114.