ORGANIC LETTERS

2013
Vol. 15, No. 6
1278–1281

Biosynthesis of 9-Methylstreptimidone Involves a New Decarboxylative Step for Polyketide Terminal Diene Formation

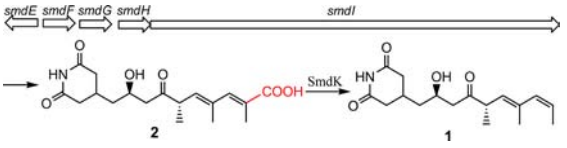
Bo Wang,[†] Yongxiang Song,[†] Minghe Luo, Qi Chen, Junying Ma, Hongbo Huang, and Jianhua Ju*

Key Laboratory of Marine Bio-resources Sustainable Utilization, Guangdong Key Laboratory of Marine Materia Medica, RNAM Center for Marine Microbiology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, 164 West Xingang Road, Guangzhou 510301, China

jju@scsio.ac.cn

Received January 25, 2013

ABSTRACT



9-Methylstreptimidone is a glutarimide antibiotic showing antiviral, antifungal, and antitumor activities. Genome scanning, bioinformatics analysis, and gene inactivation experiments reveal a gene cluster responsible for the biosynthesis of 9-methylstreptimidone in *Streptomyces himastatinicus*. The unveiled machinery features both acyltransferase- and thioesterase-less iterative use of module 5 as well as a branching module for glutarimide generation. Impressively, inactivation of *smdK* leads to a new carboxylate analogue unveiling a new mechanism for polyketide terminal diene formation.

9-Methylstreptimidone (**1**), featuring a glutarimide moiety and a terminal diene moiety, was first discovered in 1974 from *Streptomyces* sp. S885 and then rediscovered in 1989 from *Streptomyces hygroscopicus* S632.[1,2] The absolute configuration of **1** was subsequently solved by chemical and spectroscopic studies.[3] It displays antifungal,[2] antiviral, and interferon-inducing activities[4] and induces apoptosis selectively in adult T-cell leukemia cells.[5] However, thorough elucidation of the biosynthetic machinery driving production of **1** has been lacking.

Our early genome scanning strategy applied to *S. himastatinicus* ATCC 53653[6] enabled us to identify the biosynthetic gene cluster and to elucidate three cytochrome P450-catalyzed tailoring steps en route to the natural product himastatin.[7] Further analysis of genome data revealed the presence of several PKS gene clusters. Furthermore, chemical profiling of fermentation extracts reveals the presence of several peaks suggesting possible biosynthetic capabilities beyond simply those required of himastatin generation. Here we report (i) identification of 9-methylstreptimidone (**1**) and localization of a gene cluster governing its biosynthesis in *S. himastatinicus*; (ii) a model for biosynthesis of **1** based on bioinformatics analysis and gene inactivation data featuring a *trans*-acyltransferase (*trans*-AT), iterative use of module 5, a type II thioesterase (TE) that releases the polyketide chain from the assembly line; and (iii) a new SmdK-catalyzed decarboxylation process for terminal diene formation.

The fermentation broth of *S. himastatinicus*, upon LC-PDA-MS analysis, exhibited a major peak showing a quasimolecular ion peak at 308.1 $[M + H]^+$. Scaled-up

[†] These authors contributed equally.

(1) Saito, N.; Kitame, F.; Kikuchi, M.; Ishida, N. *J. Antibiot.* **1974**, *27*, 206–214.
(2) Otani, T.; Minami, Y.; Matsumoto, H.; Marunaka, T.; Lou, Z. X.; Yu, Q. W. *J. Antibiot.* **1989**, *42*, 654–661.
(3) Becker, A. M.; Rickards, R. W. *Helv. Chim. Acta* **1976**, *59*, 2393–2401.
(4) Saito, N.; Suzuki, F.; Sasaki, K.; Ishida, N. *Antimicrob. Agents Chemother.* **1976**, *10*, 14–19.
(5) Takeiri, M.; Ota, E.; Nishiyama, S.; Kiyota, H.; Umezawa, K. *Oncol. Res.* **2012**, *20*, 7–14.
(6) Kumar, Y.; Goodfellow, M. *Int. J. Syst. Evol. Microbiol.* **2008**, *58*, 1369–1378.
(7) Ma, J.; Wang, Z.; Huang, H.; Luo, M.; Zuo, D.; Wang, B.; Sun, A.; Cheng, Y.-Q.; Zhang, C.; Ju, J. *Angew. Chem., Int. Ed.* **2011**, *50*, 7797–7802.

10.1021/ol400224n
Published on Web 02/25/2013

fermentation (6 L), subsequent HPLC guided purification, and finally HRMS and ^{1}H and ^{13}C NMR data analyses revealed that the purified compound was identical to previously reported 9-methylstreptimidone (**1**).[2] Bioinformatics analysis of the genome sequence data for *S. himastatinicus* revealed that a genomic DNA segment spanning ~30 kb and consisting of 11 open reading frames (orfs) might drive the biosynthetic production of **1**. The gene cluster encodes four regulatory or resistance proteins (SmdABCD), one type II thioesterase (SmdE), one discrete AT (SmdF), one acyl carrier protein (SmdG), one amidotransferase (SmdH), one AT-less type I modular polyketide synthase (SmdI), and two decarboxylases (SmdJ, SmdK)(Table SI, Supporting Information). Beyond these, *orfs* spanning from −1 to −3 and from +1 to +3 are hypothesized to be unnecessary for the biosynthesis of **1**; this is based upon bioinformatics analysis. The sequence of the gene cluster has been deposited in the EMBL database with accession no. FR878059.

To explore the exact roles of genes implicated in the biosynthesis of **1**, each of the 17 coding genes in Table SI was inactivated by replacement with an *aac(3)IV/oriT* cassette using well-established PCR targeting method;[8] the apramycin-resistant and kanamycin-sensitive double-crossover mutants were selected and confirmed by PCR (Figures S1−S17, Supporting Information). HPLC analyses of fermentation extracts of the Δ*smdI* mutant revealed the complete absence of **1** (Figure 1), demonstrating the indispensibility of SmdI for 9-methylstreptimidone biosynthesis. Additionally, the boundaries of the gene cluster for **1** were narrowed from *SmdA* to *SmdK*; inactivations of *orf (−1)*, *orf (−2)*, *orf (−3)*, *orf (+1)*, *orf (+2)*, and *orf (+3)* had no impact on the production of **1** (Figure S19, Supporting Information).

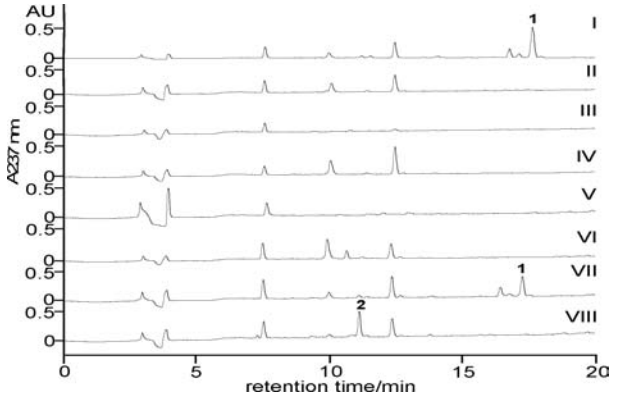


Figure 1. Metabolite profiles of *S. himastatinicus* Δ*smdEFGHIJK* mutant strains (traces II, III, IV, V, VI, VII, and VIII, respectively) in comparison with the wild-type (trace I), upon HPLC analysis. See Figure 2 for structures of **1** and **2**.

Bioinformatics suggest that *smdA−D* encode for regulatory and transportation proteins. Contrary to previous expectations, inactivation of these four genes had no influence on the production of **1** relative to that observed for wild-type *S. himastatinicus*. We further made a Δ*smd(orf(−3)-D)* mutant in which the genes from *orf (−3)* to *smdD* were all deleted and replaced with the *aac(3)IV/oriT* cassette. The resulting mutant still produced **1** with titers comparable to those of the wild-type strain. These data suggest that *smdA−D* are, in fact, not involved in the regulation of, or resistance to, **1**. Alternatively, these data suggest the remote possibility that compensatory mechanisms are available to counteract the functional loss of these *orfs*.

The *smdF* gene codes for a 294 aa *trans*-AT that shows 51%, 47%, and 48% identities to MgsB, RhiG, and OzmM from the AT-less migrastatin, rhizoxin, and oxazolomycin biosynthetic pathways, respectively.[9−11] Unlike MgsB (an AT/TE), RhiG (an AT/AT), and OzmM (an AT/Ox), all didomain *trans*-AT family proteins, SmdF is a single AT domain containing conserved amino acids for malonyl-CoA selection. Inactivation of *smdF* abolished the production of **1**, demonstrating that this *trans*-AT is crucial for 9-methylstreptimidone biosynthesis (Figure 2).

The *smdG* and *smdH* genes encode a type II acyl carrier protein (ACP) and an amidotransferase (AMT), respectively. SmdG and SmdH show high homology to MgsC (50% identity) and MgsD (57% identity), respectively, in the migrastatin biosynthetic pathway.[9] Both MgsC and MgsD have been proposed to load malonyl CoA and subsequent amino group installations prior to glutarimide formation.[9] Importantly, both Δ*smdG* and Δ*smdH* mutants failed to produce **1**, suggesting these two genes play roles similar to those of MgsC and MgsD (for migrastatin production) in the biosynthetic generation of **1** (Figure 2).

The largest gene, *smdI*, encodes a 6380 aa modular-type I PKS that can be organized into five modules. Each module minimally contains ketosynthase (KS) and acyl carrier protein (ACP) domains. Within the modules, no AT domain was found, in support of the *trans*-AT function of SmdF. All of the ACP domains possess the highly conserved signature motif DSL, in which the serine residue undergoes 4′-phosphopantetheinylation, a post-translational modification essential for polyketide biosynthesis by converting apo-ACPs into the functional holo-ACPs.[12] Functional ketoreductase (KR) domains, featuring the conserved consensus sequence GxGxxGxxA associated with NADP(H) binding,[13] are found within the 1, 3, and 5 extension modules. Functional dehydratase (DH) domains, containing the conserved consensus sequence HxxxGxxxxP,[14] are present in modules 1 and 5. Within

(8) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1541–1546.

(9) Lim, S.-K.; Ju, J.; Zazopoulos, E.; Jiang, H.; Seo, J. W.; Chen, Y.; Feng, Z.; Rajski, S. R.; Farnet, C. M.; Shen, B. *J. Biol. Chem.* **2009**, *284*, 29746–29756.

(10) Partida-Martinez, L. P.; Hertweck, C. *ChemBioChem* **2007**, *8*, 41–45.

(11) Zhao, C.; Coughlin, J. M.; Ju, J.; Zhou, D.; Wendt-Pienkowski, E.; Zhou, X.; Wang, Z.; Shen, B.; Deng, Z. *J. Biol. Chem.* **2010**, *285*, 20097–20108.

(12) Wakil, S. J. *Biochemistry* **1989**, *28*, 4523–4530.

(13) Caffrey, P. *ChemBioChem* **2003**, *4*, 654–657.

(14) Bevitt, D. J; Cortes, J.; Haydock, S. F.; Leadlay, P. F. *Eur. J. Biochem.* **1992**, *204*, 39–49.

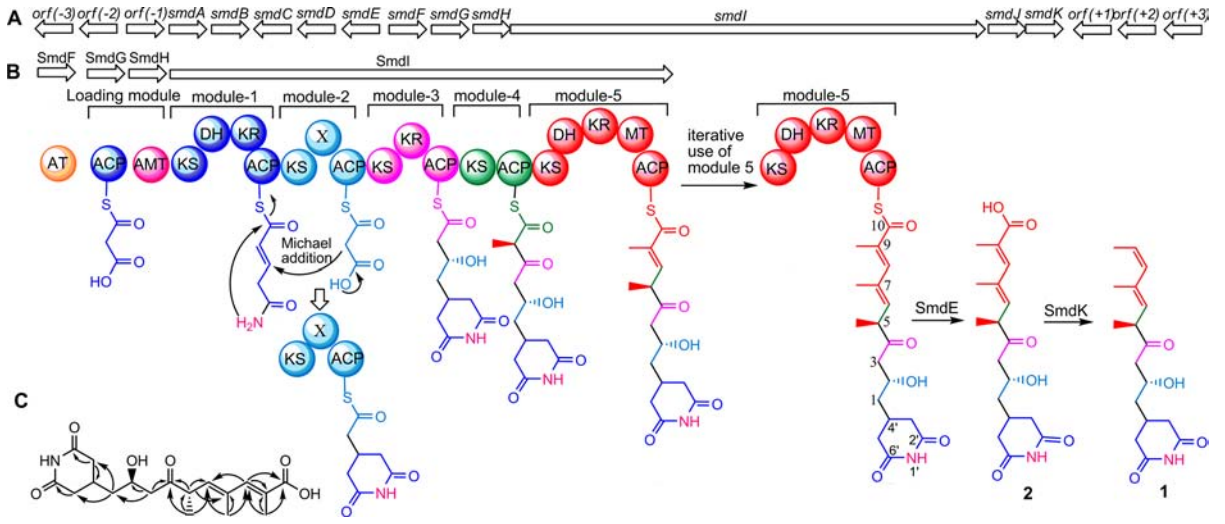


Figure 2. (A) Gene cluster organization for the 9-methylstreptimidone cluster. (B) Proposed biosynthetic pathway for **1** features AT-less, modular TE-less iterative use of module 5 and a branching domain (X) for glutarimide formation as well as a decarboxylative step for polyketide terminal diene formation as evidenced by accumulation of **2** in the Δ*smdK* mutant. (C) HMBC correlations in **2**.

module 5, a methyltransferase (MT) domain showing high homology to C-MT domains in AT-less type I PKSs, was identified. The MT domain contains a conserved glycine-rich motif I E/DxGxGxG and motif II GxxxxxxD,[15] characteristic of *S*-adenosylmethionine-dependent MTs. Since there is only one *trans*-AT protein in the gene cluster, we postulate that this single MT domain works iteratively to incorporate all three methyl groups in **1**. According to the structure of **1**, six rounds of chain elongation are required. Correspondingly, the fifth module containing KS, KR, DH, MT and ACP domains, is anticipated to be used iteratively. Surprisingly, an unknown domain (X) in module 2 was identified. A similar domain with an as yet, unknown function, has also been identified in the biosynthetic machineries for structurally related natural products migrastatin,[9] lactimidomycin, and cycloheximide (accession nos. GQ274954 and JX014302). Alignment of the X domains consisting of 470−531 aa from five other pathways reveals several motifs of conserved regions (Figure S20, Supporting Information) supporting a role in glutarimide generation. A similar domain referred to as the β-branching domain (B) in KS-B-ACP module has recently been verified in vivo in the rhizoxin pathway and is responsible for δ-lactone branch formation.[16]

We propose that the X domain plays a role in glutarimide moiety formation. X domain-catalyzed Michael addition of a malonyl unit to the double bond of the module 1-tethered intermediate with concomitant decarboxylation (Figure 2B) is likely key to glutarimide production in **1**. Following establishment of the new C−C bond that transiently links the module 1 and 2 ACPs we envision that attack of the primary amide upon the module 1 ACP thioester liberates the polyketide intermediate from module 1 and installs the glutarimide moiety (now tethered to module 2 and ready for elongation and transfer to the module 3 ACP). An alternative scenario invokes cross-modular attack of the primary amide, following Michael addition, upon the module 2 thioester. However, this proposal invokes transfer of the polyketide intermediate from the module 1 ACP directly to the module 3 ACP calling for close associations between modules 1 and 3 despite their physical separation by module 2. In the absence of data suggesting such associations and the complexity of cross modular shuffling of the polyketide intermediate called for by this mechanism, we postulate the more direct route to glutarimide formation and subsequent polyketide elaboration shown in Figure 2B.

Notably, we could find no C-terminal thioesterase (TE) domain within SmdI; such a domain had been envisioned to liberate the full-length polyketide chain once constructed by the PKS. Instead, we found that SmdE, a discrete type II TE protein with 33% identity to ChiK from *S. antibioticus* (Figure S21, Supporting Information),[17] may fulfill this role. Consistent with this hypothesis, the Δ*smdE* mutant strain failed to produce **1** (Figure 1).

Within the 9-methylstreptimidone cluster the furthest downstream *orfs* were found to be *smdJ* and *smdK* which encode two decarboxylases and are genetically cotranscribed with *smdH* and *smdI*. SmdJ shows more than 65% identity to a group of aromatic acid decarboxylases in the NCBI database. However, inactivation of *smdJ* did not alter the production of **1**, thereby excluding the involvement of its product SmdJ in 9-methylstreptimidone biosynthesis (Figure 1).

(15) Martin, J. L.; McMillan, F. M. *Curr. Opin. Struct. Biol.* **2002**, *12*, 783–793.

(16) Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. *Angew. Chem., Int. Ed.* **2009**, *48*, 5001–5004.

(17) Jia, X. Y.; Tian, Z. H.; Shao, L.; Qu, X. D.; Zhao, Q. F.; Tang, J.; Tang, G. L.; Liu, W. *Chem. Biol.* **2006**, *13*, 575–585.

BLAST analysis revealed that SmdK is a member of the UbiD superfamily of proteins. However, it shows only 32−46% identity to a variety of UbiD-like proteins from microorganisms identified on the basis of genome sequencing and whose functions have yet to be experimentally validated. Inactivation of *smdK*, affording the Δ*smdK* mutant, was found to abolish production of **1** but produced a new product peak (**2**) characterized by a quasi-molecular ion signal at m/z 350.1 $[M - H]^-$ corresponding to a carboxylated congener of **1** (Figure 1). Fermentation of the Δ*smdK* mutant (6 L scale) and isolation of **2** was carried out on a scale sufficient to provide analytically pure **2** in quantities needed for complete structure elucidation. HRMS data was consistent with a molecular formula of $C_{18}H_{25}NO_6$ for **2**, one CO_2H unit greater than that of **1**. Analysis of ^1H and ^{13}C NMR data for **2** and comparison with those of **1** disclosed that **2** bears a C-9 carboxylate (Table S6, Supporting Information). Further evidence supporting the assigned structure of **2** was obtained upon analysis of HMBC correlations (Figure 2C).

The polyketide natural products, curacin A from the marine cyanobacterium *Lyngbya majuscula* and tautomycetin from *S. griseochromogenes*, contain a terminal hydrophobic olefin or diene moiety, respectively. The former terminal alkene is formed by a TE domain with the aid of *O*-sulfation of the β-hydroxy acyl-ACP intermediate.[18] The latter is formed by the decarboxylase TtnD but also is absolutely dependent on the presence of TtnF, a dehydratase.[19] SmdK, in the biosynthetic pathway to **1**, is sufficient, by itself, to catalyze the formation of the diene moiety; this represents a new polyketide terminal decarboxylation process.

The chemical environment at C-9 in **2** bears substantial similarity to that of aromatic carboxy acids. We therefore compared the protein sequence of SmdK with those of functionally investigated aromatic acid decarboxylases. We found that SmdK displays 28.0% identity to VdcC (Q9X697) from *Streptomyces* sp. D7[20] and 34.3% identity to FDC1 (Q03034) from *Saccharomyces cerevisiae*;[21] VdcC and FDC1 are involved in vanillic acid and phenylacrylic acid nonoxidative decarboxylations, respectively. Moreover, SmdK shows 26.0% identity to Ohb1 from the anaerobe *Clostridium hydroxybenzoicum*; overexpressed Ohb1 in *E. coli* was shown to be an oxygen-sensitive, cofactor free and reversible 4-hydroxybenzoate decarboxylase.[22] Alignment of SmdK with Ohb1, FDC1, and VdcC as well as 4 other representative homologues identified by BLAST searching of the NCBI database revealed that, although these proteins share less than 46% identity with SmdK, several motifs are quite well conserved including the E-X-P motif possibly involved in substrate binding and/or catalysis (Figure 22, Supporting Information).[22]

In summary, metabolic profiling and genomic mining strategies have enabled us to, for the first time, identify a gene cluster governing the production of the glutarimide antibiotic 9-methylstreptimidone. Supported by gene inactivation experiments, an AT-less model for the biosynthesis of **1** featuring iterative use of module 5, a branching module for glutarimide formation, and a type II TE that liberates the polyketide chain from the PKS assembly line has been proposed. More importantly, a new decarboxylation process for polyketide terminal diene formation has been discovered. SmdK, showing less than 50% identity to proteins in the database, drives this unique decarboxylation process and thus serves as an exciting springboard for future mechanistic efforts.

Acknowledgment. This work was supported by MOST (2010CB833805, 2012AA092104) and the Knowledge Innovation Programs of CAS (KZCX2-YW-JC202, KSCX2-EW-G-12, and 08SL111001).

Supporting Information Available. Detailed experimental procedures, NMR data, and spectra for compounds **1** and **2**. This material is available free of charge via the Internet at http://pubs.acs.org.

(18) Gu, L.; Wang, B.; Kulkarni, A.; Gehret, J. J.; Lloyd, K. R.; Gerwick, L.; Gerwick, W. H.; Wipf, P.; Hakansson, K.; Smith, J. L.; Sherman, D. H. *J. Am. Chem. Soc.* **2009**, *131*, 16033–16035.

(19) Luo, Y.; Li, W.; Ju, J.; Yuan, Q.; Peters, N. R.; Hoffmann, F. M.; Huang, S.-X.; Bugni, T. S.; Rajski, S.; Osada, H.; Shen, B. *J. Am. Chem. Soc.* **2010**, *132*, 6663–6671.

(20) Chow, K. T.; Pope, M. K.; Davies, J. *Microbiology*. **1999**, *145*, 2393–2403.

(21) Mukai, N.; Masaki, K.; Fujii, T.; Kawamukai, M.; Iefuji, H. *J. Biosci. Bioeng*. **2010**, *109*, 564–569.

(22) Huang, J.; He, Z.; Wiegel, J. *J. Bacteriol*. **1999**, *181*, 5119–5122.

The authors declare no competing financial interest.