Organic LETTERS

Letter

pubs.acs.org/OrgLett

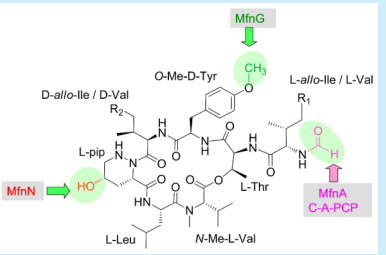
Biosynthesis of the Anti-infective Marformycins Featuring Pre-NRPS Assembly Line *N*-Formylation and *O*-Methylation and Post-Assembly Line *C*-Hydroxylation Chemistries

Jing Liu,[†] Bo Wang,[†] Hongzhi Li, Yunchang Xie, Qinglian Li, Xiangjing Qin, Xing Zhang, and Jianhua Ju*

CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, 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

Supporting Information

ABSTRACT: The biosynthetic gene cluster governing production of anti-infective marformycins was identified from deep sea-derived *Streptomyces drozdowiczii* SCSIO 10141. The putative *mfn* gene cluster (45 kb, 20 orfs) was found to encode six NRPSs and related proteins for cyclodepsipeptide core construction (*mfnCDEFKL*), a methionyl-tRNA formyltransferase (*mfnA*), a SAM-dependent methyltransferase (*mfnG*), and a cytochrome P450 monooxygenase for piperazic acid moiety hydroxylation (*mfnN*); notably, only MfnN uses an intact cyclodepsipeptide intermediate as its substrate.



Cyclic peptides have played an important role in natural product drug discovery by virtue of their structural diversity, well understood biosynthetic pathways, cell permeability, and resistance to proteolytic digestion/inactivation pathways.[1−4] Many cyclopeptides such as the immunosuppressive cyclosporine A, antibacterial agents tyrocidine A, colistin, capreomycin and daptomycin, antifungal agent caspofungin, and antitumor agent romidepsin are commonly used clinical drugs. Two marine-derived cyclodepsipeptides, aplidine (plitidepsin or dihydrodidemnin B), originally isolated from the Mediterranean tunicate *Aplidium albicans*,[5,6] and kahalalide F (PM-92012), originally isolated from the sacoglossan mollusk *Elysia rufescens* and then subsequently from the algae *Bryopsis pennata* upon which it feeds,[7,8] are currently being assessed as anticancer agents in phase II clinical trials by PharmaMar.[9,10]

We recently isolated a group of cyclodepsipetides, marformycins A−F (MFN, **1**−**6**), from fermentations of a deep sea sediment-derived *Streptomyces drozdowiczii* SCSIO 10141. The structures of **1**−**6**, including their absolute stereochemistry, have been determined by chiral-phase HPLC and single-crystal X-ray diffraction data analyses by our group; all have been found to consist of five nonproteinogenic amino acid residues implicating the importance of nonribosomal peptide synthetases (NRPSs) in their construction. These residues include a piperazic acid unit, an *O*-methyl-D-Tyr, a D-*allo*-Ile/D-Val, an L-*allo*-Ile/L-Val, and an *N*-methyl-Val.[11] These compounds display no cytotoxicity, appearing, instead, to exert selective anti-infective activity against *Micrococcus luteus* and the bacteria *Propionibacterium acnes* and *P. granulosum*, highlighting their potential as anti-infective drug leads (e.g., for treating acne vulgaris).[11,12] To delineate the biosynthetic machinery of this scaffold and to enable engineering of new analogues, we report herein: (i) identification of the gene cluster governing MFN biosynthesis in *Streptomyces drozdowiczii* SCSIO 10141; (ii) *in vivo* genetic and *in vitro* biochemical characterization of *mfnG* as both an L-and D-tyrosine-*O*-methyltransferase; and (iii) *in vivo* inactivation of *mfnA*, *mfnK* *mfnG*, and *mfnN*, coupled with bioinformatics analysis and metabolite identification revealing that the MFNs result from the *N*-terminus preassembly line formylation and pre-NRPS assembly line *O*-methylation of the Tyr residue as well as a regio- and stereoselective post-NRPS hydroxylation of the piperazic acid residue.

That the MFNs contain nonproteinogenic amino acid units supported our initial hypothesis that MFN construction is governed, in large part, by NRPSs. Accordingly, we sought to identify the *mfn* cluster by application of whole genome scanning of *Streptomyces drozdowiczii* SCSIO 10141 using 454 sequencing technologies. Upon genome data annotation, we localized a 45 kb DNA segment consisting of 20 opening reading frames (orfs) likely to be involved in MFN biosynthesis. The putative *mfn* gene cluster (Supporting Information, Table S1) was found to encode (i) six NRPSs and related proteins for cyclodepsipeptide core construction (MfnCDEFKL), (ii) a methionyl-tRNA formyltransferase (MfnA), a SAM-dependent methyltransferase (MfnG), and a cytochrome P450 monooxygenase (MfnN) for scaffold tailoring steps, (iii) a regulatory protein (MfnM) and a transportation protein (MfnR), and (iv) seven proteins associated with precursor supply or having unclear biosynthetic function (MfnBHIJOPQ). The nucleotide sequences have been deposited in the GenBank with accession number KP715145, and the genetic organization of the biosynthetic gene cluster is shown in Figure 1.

Received: February 6, 2015

DOI: 10.1021/acs.orglett.5b00389
Org. Lett. XXXX, XXX, XXX−XXX

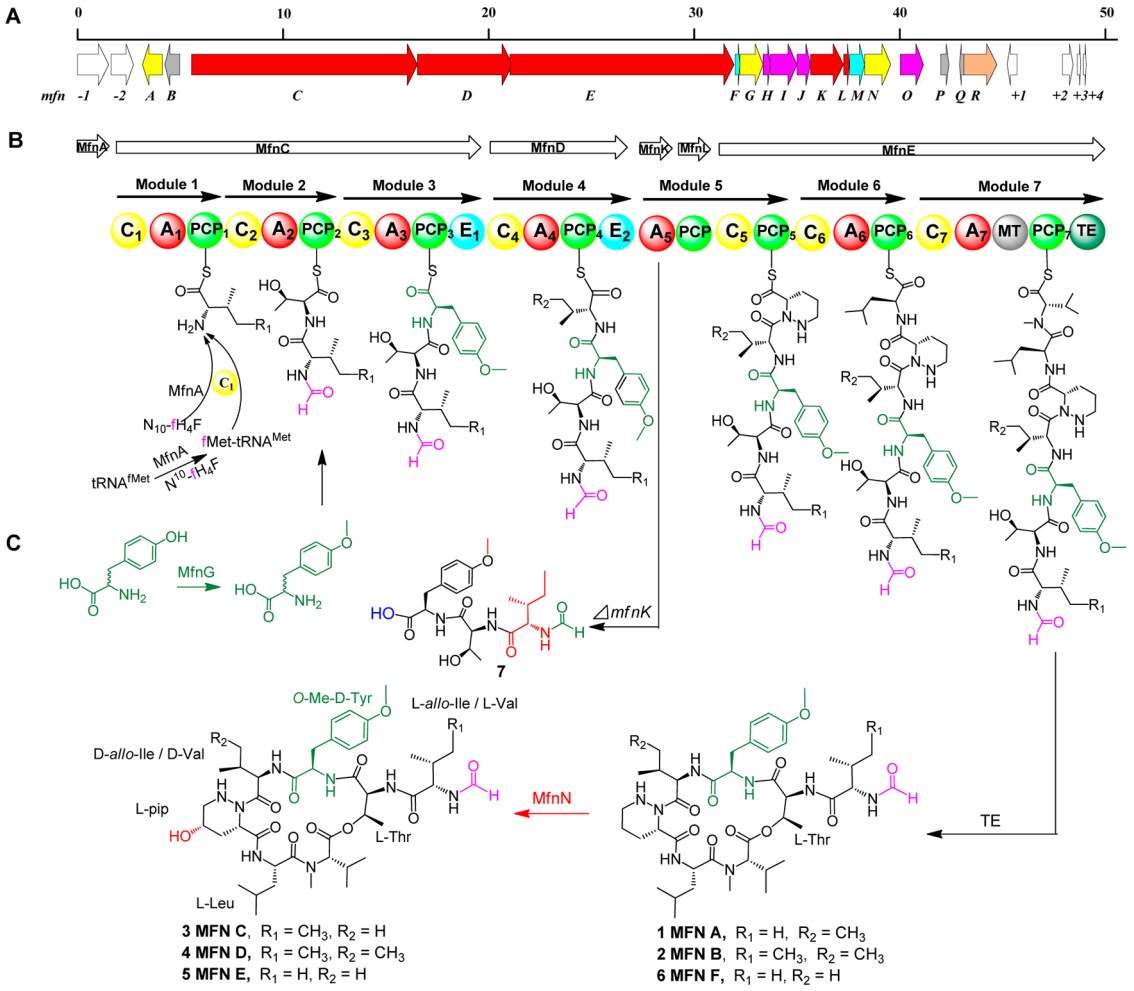


Figure 1. (A) Organization of *mfn* biosynthetic gene cluster; (B) proposed biosynthetic pathway for the MFNs; (C) MfnG-catalyzed conversion of L(D)-Tyr into *O*-methyl-L(D)-Tyr. *N*-Formylated intermediate **7** was identified from the engineered Δ*mfnK* mutant. N^{10}-fH$_4$F: N^{10}-formyltetrahydrofolate; fMet-tRNA^{Met}: formylmethionine initiator tRNA.

We constructed a genomic library of *S. drozdowiczii* SCSIO 10141 using the SuperCos1 vector system and screened it with selected PCR primers to yield 16 positive clones. The boundary of the *mfn* gene cluster was determined by bioinformatics analysis and PCR-targeting mutagenesis methods that involved replacing each targeted gene with an spectinomycin resistance cassette.[13,14] On the basis of these efforts, it was found that *orf(−1)* encodes a highly conserved DNA recombination−mediator protein and that *orf(−2)* encodes a hypothetical protein; disruptions of these two flanking genes *orf(−1)* and *orf(−2)* did not affect MFN production, indicating their dispensability for MFN biosynthesis. Inactivations of *mfnB* (encodes a LipE-like protein) and *mfnP* and *mfnQ* (both encode hypothetical proteins) also did not affect the MFN production, suggesting they are not necessary in MFN biosynthesis (Supporting Information, Figure S1).

We then analyzed genes anticipated to be critical to MFN core construction. Three NRPS proteins MfnC, MfnD, MfnE, a free adenylation (A) enzyme MfnK, and a free peptidyl carrier protein (PCP) MfnL constituting a total of seven modules were identified and envisioned to drive generation of the MFN core scaffold (Figure 1). The predicted specificity of A domains for amino acid activation (Supporting Information, Table S7) suggested that MFN core peptide synthesis starts with L-Val/L-*allo*-Ile as dictated by MfnC module 1 and terminates with the addition of *N*-Me-L-Val carried out by MfnE module 7. This general paradigm agrees well with the pattern of amino acid residues found in the MFN natural products. Inactivation of *mfnD* abolished the MFN production (Figure 2, trace ii), demonstrating its involvement in the MFN core construction.

Besides the three basic domains that constitute a minimal NRPS [i.e., the condensation (C), A, and PCP], within the putative *mfn* cluster there are two epimerization (E) domains (modules 3 and 4) and a methyltransferase domain (MT) in module 7, consistent with the presence of D-Tyr, D-Val/D-*allo*-Ile, and *N*-Me-L-Val units in the MFN scaffold. Additionally, the *mfnF* gene encodes a putative MbtH-like protein composed of only 76 amino acids. Recently, it was discovered that these MbtH-like proteins are required for selected adenylation reactions in antibiotic pathways; MbtH-like protein inactivation has been correlated to dramatically reduced antibiotic yields.[15] Indeed, the Δ*mfnF* mutant strain generated herein was found incapable of MFN production (Figure 2, trace iii), consistent with previous reports and suggesting an essential role in MFN NRPS assembly line function.

Upstream of the MFN cluster resides *mfnA* encoding a methionyl-tRNA (Met-tRNA^{Met}) formyltransferase, containing a typical N^{10}-formyltetrahydrofolate (N^{10}-fH$_4$F) binding motif (Figure S2, Supporting Information). In considering the structure of MFNs, we envisioned *mfnA* to be involved in *N*-formylation of the MFNs. We inactivated MfnA using PCR-

DOI: 10.1021/acs.orglett.5b00389
Org. Lett. XXXX, XXX, XXX−XXX

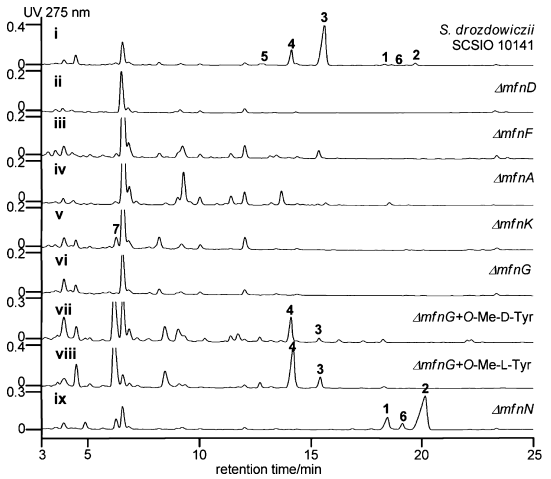


Figure 2. HPLC profiles of *S. drozdowiczii* SCSIO 10141 wild-type and mutant strains, Δ*mfnG* mutant provided with *O*-Me-Tyr (L or D isomers) in the fermentation. See Figure 1 for structures.

targeting mutagenesis methodology directed at *mfnA*; fermentation and HPLC-based extract analysis for the Δ*mfnA* strain revealed that MfnA inactivation abolished production of MFN and corresponding intermediates (Figure 2, trace iv). These data clearly demonstrate the indispensability of MfnA to MFN production and suggest that it may direct incorporation of the formyl group as a starter unit in the L-Val/L-*allo*-Val lateral chain (Figure 1).

In considering MfnA and its precise role in MFN synthesis, we noted that MfnC harbors a noncanonical C-A-PCP initiation module (Figure 1); this unusual starter module has been discovered in assorted *N*-acylated NRPS products such as the xenematides.[16] We thus proposed that MfnA may transfer a formyl group to Met-tRNA^{Met} from the primary metabolite pathway to generate formylmethionine tRNA (fMet-tRNA^{Met}). The C-domain of module 1 (also containing the conserved HHxxxD motif, Supporting Information, Figure S3) would condense the fMet-tRNA^{Met} with L-Val/L-*allo*-Val that had been loaded onto the module 1 PCP, thus forming *N*-formyl L-Val/L-*allo*-Val of the MFN biosynthetic assembly line. Aminoacylated tRNA-dependent peptide bond formation mediated by PacB, but not catalyzed by the condensation domain, has been validated in the nucleoside antibiotic pacidamycin pathway.[17] Alternatively, MfnA may directly transfer the formyl group from N^{10}-fH$_4$F to the L-Val on the PCP domain of module 1, bypassing the C domain of module 1. Notably, a formylation (F) domain embedded within the NRPS module (F-A-PCP) has been shown to transfer the *N*-formyl group from N^{10}-fH$_4$F to a PCP-tethered L-Val in the gramicidin pathway.[18]

A pre-NRPS assembly line *N*-formylation mechanism is supported by LC−MS-based metabolite analysis of fermentations of a Δ*mfnK* mutant which revealed the accumulation of small quantities of a biosynthetic intermediate **7** (Figure 2, trace v). The compound **7** was structurally characterized following large-scale fermentation, purification, and subsequent MS and ^{1}H and ^{13}C NMR data analyses. It is notable that, despite its small size, compound **7** bears terminal *N*-formylation indicative of MfnA action at the start of MFN assembly.

We then explored the role of *mfnG* encoding a putative methyltransferase. MfnG shows high similarity (68%) to the *S*-adenosylmethionine (SAM)-dependent *O*-methyltransferase Sky37 from *Streptomyces* sp. acta 2897 and was thus proposed to carry out Tyr *O*-methylation during MFN assembly. However, *mfnG* inactivation completely ablated MFN production (Figure 2, trace vi). This suggests that MfnG may methylate free Tyr; the resulting *O*-methyl-Tyr may be the species recognized by the A$_3$ domain and subsequently activated enabling introduction to the NRPS assembly line. To verify this hypothesis, we added *O*-methyl-D-Tyr and *O*-methyl-L-Tyr precursors, respectively, to fermentations of the Δ*mfnG* mutant strain. As expected, the ability of the Δ*mfnG* strain to produce MFNs was restored by the addition of *O*-methyl-L-Tyr to the fermentation medium. However, a substantial amount of MFN production was also observed with the addition of *O*-methyl-D-Tyr. The ratio of harvested MFNs obtained by providing Δ*mfnG* access to both *O*-methyl-L-Tyr and *O*-methyl-D-Tyr was approximately 3:1 (Figure 2, traces vii and viii). These data suggest that the A$_3$ domain of MfnC possesses relaxed substrate specificity but prefers *O*-methyl-L-Tyr over *O*-methyl-D-Tyr.

The biochemical activity of MfnG was further explored *in vitro*. The *mfnG* gene was cloned into the *Nde*I and *Eco*RI sites of the pET28a (+) vector, overexpressed in *Escherichia coli* BL21 (DE3), and purified to homogeneity as an *N*-terminus His-tagged protein. The enzymatic activity of MfnG in various buffer systems (i.e., HEPES, phosphate, and Tris buffers) and over a wide pH range (5.0−9.0) was evaluated. The optimal enzyme activity was observed in a pH 8.0 phosphate buffer system at 30 °C. Typical *in vitro* assays (50 μL) were performed in the presence of 50 mM phosphate buffer (pH 8.0), 0.2 mM substrate L-Tyr or D-Tyr, 0.24 mM SAM, and 0.5 μM MfnG at 30 °C for 2 h, and were quenched by addition of two volumes of MeOH. HPLC data were acquired at different time points revealing MfnG-catalyzed generation of *O*-methyl-L-Tyr and *O*-methyl-D-Tyr from L-Tyr and D-Tyr, respectively, in a time-dependent manner (Figure 3). Detailed

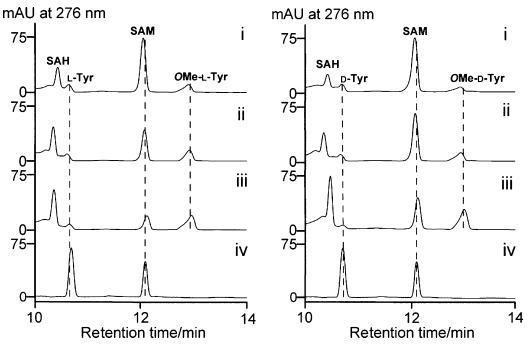


Figure 3. HPLC analyses MfnG-catalyzed reactions using substrates L-Tyr and D-Tyr. Time-course experiments incubated with MfnG and SAM for (i) 10 min, (ii) 30 min, and (iii) 60 min; panel (iv) in each set of analyses represents a control reaction in which MfnG was heat denatured prior to substrate presentation.

kinetic analyses of MfnG were carried out using a 50 μL reaction system containing 50 mM sodium phosphate buffer (pH 8.0). The kinetic parameters of MfnG-catalyzed reactions with L- and D-Tyr were determined and are summarized in Table 1; these kinetic data reveal that MfnG utilizes L-Tyr as the preferred substrate in accord with the previous feeding results.

Finally, we examined the role of putative cytochrome P450 enzyme MfnN. MfnN displays high similarity (77%) to and identity (64%) with Ema5 from the abamectin biosynthetic pathway; Ema5 converts avermectin to 4″-oxoavermectin.[19] The *mfnN* gene was inactivated in a manner similar to that used to inactivate *mfnG*. HPLC analyses of culture extracts revealed that

DOI: 10.1021/acs.orglett.5b00389
Org. Lett. XXXX, XXX, XXX−XXX

Table 1. Kinetic Parameters of MfnG-Catalyzed Reactions

substrate	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (μM^{-1} min^{-1})
L-Tyr	78.74 ± 7.790	22.73 ± 0.5665	2.87 ± 10^{-1}
D-Tyr	173.8 ± 13.90	22.07 ± 0.8226	1.27 ± 10^{-1}

the Δ*mfnN* mutant failed to produce MFNs C (**3**) and D (**4**) but generated MFNs A (**1**) and B (**2**) as the predominant products (Figure 2, trace ix). The structures of **1** and **2** produced by the Δ*mfnN* mutant were found to be identical to MFNs A and B isolated from the wild-type strain upon large scale (8-L) fermentation, purification, and HRMS, ^{1}H NMR, and ^{13}C NMR data comparisons. These data clearly reveal that MfnN is a cytochrome P450 monooxygenase responsible for the regio- and stereospecific hydroxylation of the MFN piperazic acid moiety.

MFNs A–F (**1**–**6**) belong to a very small class of *N*-terminally formylated nonribosomal peptide natural products. Members of this family include the linear gramicidins,[18] cyclic anabaenopeptilides,[20] and the siderophores rhodochelin and coelichelin.[21] In the rhodochelin biosynthetic pathway, installation of the *N*-formyl group is carried out by the formyltransferase Rft using N^{10}-fH_4F and the small molecule *δ*-*N*-hydroxyornithine as substrates.[21] In considering the MFN machinery, the noncanonical initiation module with a C-A-PCP arrangement, the critical role of the methionyl-tRNA formyltransferase MfnA and the captured NRPS off-line *N*-formylated biosynthetic intermediate **7** strongly suggests a different paradigm for *N*-terminus formylation of NRPS natural products. Specifically, the *mfn* cluster indicates the possibility that formylation may be accomplished via condensation domain-mediated peptide bond formation using fMet-$tRNA^{Met}$ and the PCP-tethered L-Val residue as substrates. Alternatively, MfnA (not the F domain)-mediated peptide bond formation may take place using N^{10}-fH_4F and the PCP-tethered L-Val residue as substrates.

Among NRPS-based systems, it is a unique feature of the *mfn* machinery that preassembly line methylation of Tyr takes place. More surprising is that the agent of this change, MfnG, can use both L- and D-Tyr as substrates yet the natural MFNs contain exclusively the *O*-methyl-D-Tyr moiety. These observations are especially intriguing in light of the putative epimerase domains embedded within the NRPS machinery. Moreover, these findings provide clues relevant to understanding the biosynthetic pathway of the cyclic peptide aplidine. Importantly, aplidine, now in clinical trials, also contains a methyl-*O*-Tyr moiety whose generation remains unclear.[5]

In summary, we have employed genome scanning, bioinformatics analyses, gene inactivations, and *in vitro* biochemical experiments to identify and validate the 45 kb *mfn* biosynthetic gene cluster in *Streptomyces drozdowiczii* SCSIO 10141. Significant attention has been paid to tailoring steps involved in MFN production. These steps include two unusual pre-NRPS assembly line transformations, (i) an *N*-terminal formylation step and (ii) a pre-NRPS *O*-methylation step, to afford *O*-methyl-L-Tyr and *O*-methyl-D-Tyr as MFN building blocks (although only the D isomer appears within completed MFN scaffolds). Following completion of NRPS-directed steps and liberation from the NRPS machinery, a novel, regio- and stereospecific hydroxylation of the MFN piperazic acid moiety was carried out (by MfnN). These findings set the stage for full characterization of the biosynthetic mechanisms driving MFN production and enable the engineering of new MFN analogues using combinatorial or precursor-directed biosynthetic strategies for antibacterial drug discovery.

■ ASSOCIATED CONTENT

Supporting Information

Detailed experimental procedures, NMR data and spectra for compound **7**. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jju@scsio.ac.cn.

Author Contributions

†J.L. and B.W. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by NSFC (41306146, 31290233, 81425022), and MOST (2012AA092104).

■ REFERENCES

(1) Andavan, G. S.; Lemmens-Gruber, R. *Mar. Drugs.* **2010**, *8*, 810–834.

(2) Lemmens-Gruber, R.; Kamyar, M. R.; Dornetshuber, R. *Curr. Med. Chem.* **2009**, *16*, 1122–1137.

(3) Sarabia, F.; Chammaa, S.; Ruiz, A. S.; Ortiz, L. M.; Herrera, F. J. *Curr. Med. Chem.* **2004**, *11*, 1309–1332.

(4) Ballard, C. E.; Yu, H.; Wang, B. *Curr. Med. Chem.* **2002**, *9*, 471–498.

(5) Schmitz, F. J.; Yasumoto, T. *J. Nat. Prod.* **1991**, *54*, 1469–1490.

(6) Xu, Y.; Kersten, R. D.; Nam, S. J.; Lu, L.; Al-Suwailem, A. M.; Zheng, H.; Fenical, W.; Dorrestin, P. C.; Moore, B. S.; Qian, P. Y. *J. Am. Chem. Soc.* **2012**, *134*, 8625–8632.

(7) Hamann, M. T.; Scheuer, P. J. *J. Am. Chem. Soc.* **1993**, *115*, 5825–5826.

(8) Ashour, M.; Edrada, R.; Ebel, R.; Wray, V.; Wätjen, W.; Padmakumar, K.; Müller, W. E.; Lin, W. H.; Proksh, P. *J. Nat. Prod.* **2006**, *69*, 1547–1553.

(9) Mayer, A. M.; Glaser, K. B.; Cuevas, C.; Jacobs, R. S.; Kem, W.; Little, R. D.; McIntosh, J. M.; Newman, D. J.; Potts, B. C.; Shuster, D. E. *Trends Pharmacol. Sci.* **2010**, *31*, 255–265.

(10) Newman, D. J.; Cragg, G. M. *Mar. Drugs* **2014**, *12*, 255–278.

(11) Zhou, X.; Huang, H.; Li, J.; Song, Y.; Jiang, R.; Liu, J.; Zhang, S.; Hua, Y.; Ju, J. *Tetrehedron* **2014**, *70*, 7795–7801.

(12) Watanabe, M.; Kamigiri, K.; Tanaka, K.; Takeda, Y.; Yokoi, T.; Shibazaki, K.; Suzumura, K.; Hashimoto, M.; Nishiwaki, S.; Takase, S.; Abe, F. PCT Patent WO 2006073151, 2006.

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

(14) Zhang, Y.; Huang, H.; Chen, Q.; Luo, M.; Sun, A.; Song, Y.; Ma, J.; Ju, J. *Org. Lett.* **2013**, *15*, 3254–3257.

(15) Zhang, W.; Heemstra, J. R., Jr.; Walsh, C. T.; Imker, H. J. *Biochemistry* **2010**, *49*, 9946–9947.

(16) Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. *Org. Lett.* **2011**, *13*, 5144–5147.

(17) Zhang, W.; Ntai, I.; Kelleher, N. L.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 12249–12253.

(18) Schoenafinger, G.; Schracke, N.; Linne, U.; Marahiel, M. A. *J. Am. Soc. Chem.* **2006**, *128*, 7406–7407.

(19) Jungmann, V.; Molnár, I.; Hammer, P. E.; Hill, D. S.; Zirkle, R.; Buckel, T. G.; Buckel, D.; Ligon, J. M.; Pachlatko, J. P. *Appl. Environ. Microbiol.* **2005**, *71*, 6968–6976.

(20) Rouhiainen, L.; Paulin, L.; Suomalainen, S.; Hyytiäinen, H.; Buikema, W.; Haselkorn, R.; Sivonen, K. *Mol. Microbiol.* **2000**, *37*, 156–167.

(21) Bosello, M.; Mielcarek, A.; Giessen, T. W.; Marahiel, M. A. *Biochemistry* **2012**, *51*, 3059–3066.

DOI: 10.1021/acs.orglett.5b00389
Org. Lett. XXXX, XXX, XXX–XXX