JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

The Fumagillin Biosynthetic Gene Cluster in *Aspergillus fumigatus* Encodes a Cryptic Terpene Cyclase Involved in the Formation of β-*trans*-Bergamotene

Hsiao-Ching Lin,[†] Yit-Heng Chooi,[†] Sourabh Dhingra,[‡] Wei Xu,[†] Ana M. Calvo,[‡] and Yi Tang*[,†,§]

[†]Department of Chemical and Biomolecular Engineering and [§]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States

[‡]Department of Biological Sciences, Northern Illinois University, DeKalb, Illinois 60115, United States

Supporting Information

ABSTRACT: Fumagillin **1** is a meroterpenoid from *Aspergillus fumigatus* that is known for its anti-angiogenic activity by binding to human methionine aminopeptidase 2. The genetic and molecular basis for biosynthesis of **1** had been an enigma despite the availability of the *A. fumigatus* genome sequence. Here, we report the identification and verification of the *fma* gene cluster, followed by characterization of the polyketide synthase and acyltransferase involved in biosynthesis of the dioic acid portion of **1**. More significantly, we uncovered the elusive β-*trans*-bergamotene synthase in *A. fumigatus* as a membrane-bound terpene cyclase.

Modern genome sequencing technologies have significantly improved our abilities in identifying biosynthetic gene clusters of natural products. However, gene cluster assignment relying on homology search of enzymes catalyzing key biosynthetic transformations can sometimes fail, largely due to nature's flexibility in using seemingly unrelated enzymes to catalyze these reactions. In these cases, the link between the natural product and the gene cluster cannot be easily established despite a sequenced and annotated genome. One example is the meroterpenoid fumagillin **1** produced by the pathogenic fungus *Aspergillus fumigatus*. Since its discovery in the 1950s,[1] **1** and related compounds have been intensely studied because of their potential use in treating amebiasis[2] and microsporidiosis,[3] and recently for their anti-angiogenic properties from the irreversible inhibition of human type 2 methionine aminopeptidase (MetAP2).[4,5] Structurally, **1** consists of a highly oxygenated cyclohexane-containing terpenoid (fumagillol, **2**) that is esterified with a decatetraenedioic (dioic) acid.[6] The intriguing structure of **1** has served as a total synthesis target for the past 40 years.[7] **2** has also been remodeled to yield new structures via a reaction discovery approach.[8] Semisynthetic analogues of **1** with modified ester group have undergone clinical trials for treating different cancers.[9] Hence, elucidating the biosynthesis of **1** can have important biomedical ramifications.

Based on the feeding study of a structurally related compound ovalicin (Figure 1),[10] **2** is likely derived from the sesquiterpene β-*trans*-bergamotene **3**, which can also be isolated from *A. fumigatus*.[11] The enzymatic mechanism for the formation of **3** from farnesyl-PP (FPP) has been previously studied using cell-free extracts from the ovalicin-producing fungus *Pseudeurotium ovalis*,[12,13] while the rearrangement of **3** to ovalicin have been probed by isotope incorporation studies in the context of ovalicin biosynthesis.[14−17] Nevertheless, the genetic and molecular bases for the biosynthesis of **3**, **2** and ultimately **1** have remained unknown. While genome sequencing of *A. fumigatus* Af293 provided initial hope that the gene cluster of **1** can be located, it was soon realized that there is no terpene cyclase homologue present in the entire genome.[18] Because of this highly unexpected finding, the gene cluster has not been identified 8 years after the sequencing report. More importantly, the lack of a terpene cyclase candidate in *A. fumigatus* suggests that the cyclization of FPP into **3** must be catalyzed by a noncanonical terpene cyclase that is concealed from common homology search strategies. Cryptic terpene cyclases in other fungal meroterpenoid gene cluster have been noted.[19] Here we describe the verification of the gene cluster of **1** and identification of an integral membrane protein as the terpene cyclase responsible for the formation of **3** from FPP.

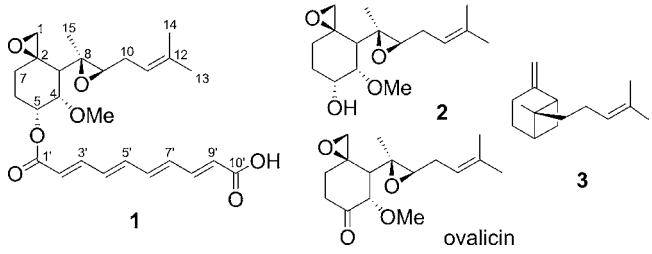

Figure 1. Fumagillin **1** and related compounds.

Since the dioic acid moiety in **1** mostly likely originates from a reduced polyketide pathway, we scanned all the highly reducing PKS (HR-PKS) genes in the *A. fumigatus* genome and their neighboring genes. Our attention was drawn to a cluster (hereby named *fma*) containing HR-PKS gene AFUA_8G00370 (Af370) (Table S1). Most notably, the cluster encodes a MetAP-2 (AFUA_8G00410 or Af410) and a type 1 MetAP. Since **1** is known to inhibit eukaryotic MetAP-2,[4] we postulated that Af410 may encode a self-resistant enzyme for *A. fumigatus* toward **1**. Indeed, an additional housekeeping MetAP-2 (AFUA_2G01750) was found in *A. fumigatus,* which is well-

Received: December 21, 2012
Published: March 14, 2013

   dx.doi.org/10.1021/ja312503y | *J. Am. Chem. Soc.* 2013, 135, 4616−4619

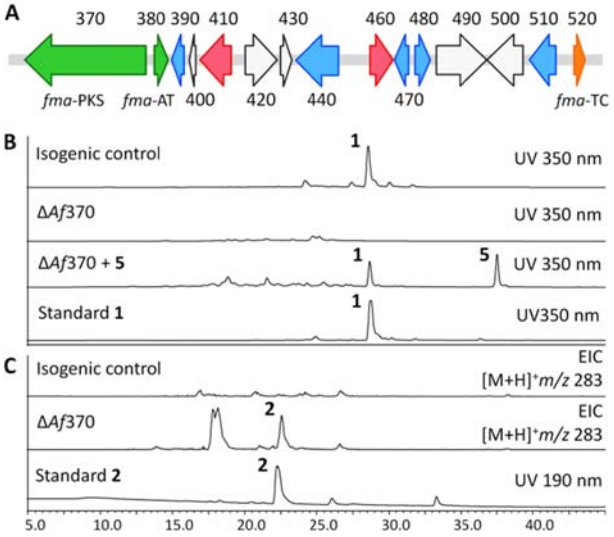

Figure 2. Genetic verification of *fma*-PKS in the biosynthesis of **1**. (A) The *fma* gene cluster. The numbers are part of the locus tag AFUA_8G00###, e.g., 370 is AFUA_8G00370. Genes are color-coded for polyketide biosynthesis (green), terpene cyclization (orange), tailoring modification (blue), self-resistance gene (red), and unknown/unassigned function (white) (see Table S1). (B) HPLC analysis of metabolites extracted from wild type and ΔAf370 strain showing loss of **1** and restored production of **1** upon feeding of **5** to ΔAf370. (C) LC-MS analysis of the ΔAf370 strain showing the accumulation of **2**.

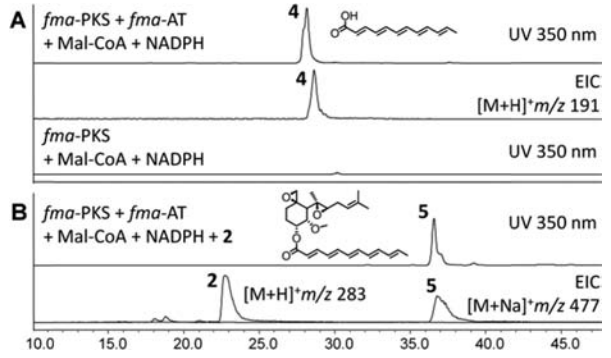

Figure 3. In vitro verification of the functions of *fma*-PKS and *fma*-AT in the biosynthesis of **4** and **5**. (A) *fma*-PKS synthesized the C_{12} polyketide product that is hydrolyzed to yield **4** by *fma*-AT. (B) Acylation of **2** (1 mM) yielded **5** by *fma*-PKS and -AT.

conserved among all sequenced aspergilli. A number of genes in the vicinity of the Af370 and Af410 also encode enzymes that are consistent with the structure of **1** and **2**, including four oxidases (to initiate oxidative rearrangement of **3**, to introduce the bis-epoxide, to insert the C4,C5-*syn*-diol, and to oxidize the polyketide methyl terminus); a methyltransferase that may *O*-methylate C4-OH; and an α,β-hydrolase that may release or transfer the polyketide from HR-PKS to **2** (Table S1). Thus, despite the absence of a canonical terpene cyclase, we hypothesized that the *fma* genes in Table S1 likely comprise the target cluster.

We first deleted Af370 encoding the HR-PKS (named *fma*-PKS) in *A. fumigatus* CEA17 *akuB*^{KU80} strain (*pyrG89*, Δ*akuB*^{KU80}), which is deficient in non-homologous end joining.[20] In comparison to the isogenic control strain, the metabolic profile of ΔAf370 analyzed by LC-MS showed the complete abolishment of **1** (Figure 2B) and the accumulation of **2** (Figure 2C), which is barely detectable in the isogenic strain. To further prove the role of *fma*-PKS, the entire 264 kDa enzyme was expressed from *Saccharomyces cerevisiae* BJ5464-NpgA, which has been used for soluble expression of large fungal PKSs.[21] The 6xHis-tagged enzyme was purified by Ni^{2+}-affinity chromatography (Figure S2A). In vitro reaction with malonyl-CoA and NADPH resulted in a yellow product that cannot be extracted using ethyl acetate with 1% acetic acid (Figure S2B). This result hinted that a conjugated polyunsaturated fatty acyl product may be synthesized and remain attached to the *fma*-PKS in the absence of a releasing enzyme. Since the product of *fma*-PKS is to be transacylated to C5-hydroxyl of **2**, we hypothesized that a dedicated acyltransferase may be involved.

In the *fma* gene cluster, a putative α/β hydrolase is encoded immediately downstream of Af370 by AFUA_8G00380 (Af380), which could be involved in PKS product release. Alternatively, the polyketide chain may also be transferred directly to **2** by an acyltransfease, such as LovD and CazE involved in the biosynthesis of lovastatin[22] and chaetoviridin,[23] respectively. To examine the role of the enzyme encoded by Af380 (named *fma*-AT), the 6xHis-tagged enzyme was solubly expressed from BJ5464-NpgA and purified to homogeneity (Figure S2A). Coincubation of *fma*-PKS and -AT with malonyl-CoA and NADPH resulted in the formation of a single yellow compound **4** that can be extracted into the organic phase (Figure 3A). LC-MS analysis of **4** yielded a mass of 190.1 ($C_{12}H_{14}O_2$), consistent with that of dodecapentaenoic acid (Figure 3A). Repeating the assay with [2-^{13}C]-malonate and a MatB-dependent *in situ* malonyl-CoA regeneration system resulted in an increase of the *m*/*z* of **4** by 6 mu, which is a further confirmation of **4** as a hexaketide. Although **4** can be detected in the culture of BJ5464-NpgA co-expressing *fma*-PKS and -AT (Figure S4), the compound readily degraded and is unstable upon purification.

To examine if *fma*-AT can transfer the polyketide product of *fma*-PKS to **2**, we performed the *fma*-PKS turnover assay in the presence of *fma*-AT and **2**. Interestingly, **2** can be converted to a single new product **5** with the mass of 454.3 (*m*/*z* 477.1 [M +Na]$^+$), corresponding to that of **2** acylated with **4**. To verify the structure of **5**, 2.5 mg of **2** was supplemented to a 50 mL yeast culture co-expressing *fma*-PKS and -AT (Supporting Information (SI)). Following 2 days of culturing, 1.1 mg of **5** can be purified from the culture, which allowed complete NMR characterization of **5** by 1D and 2D NMR, and confirmed the structure to be that shown in Figure 3B. To verify **5** is a true on-pathway intermediate in the biosynthesis of **1**, ~40 μg/mL of **5** was supplied to the culture of ΔAf370 strain. Following 2 days of culturing, restoration of the biosynthesis of **1** was observed (Figure 2B). Collectively, our genetic and biochemical studies on *fma*-PKS and -AT confirm the involvement of these enzymes in biosynthesis of **1**, and provide the first evidence for this being the authentic *fma* gene cluster.

In light of the gene cluster identification and above biochemical findings, a putative pathway for **1** is proposed as shown in Figure 4. The terpenoid origin of the carbon backbone of **1** via **3** has been previously established.[6,10,11] **3** is cyclized from FPP via nerolidyl diphosphate (NPP) followed by a bisabolyl cation intermediate.[12,13] However, the previously proposed intermediate **9** (Figure 4, route 1)[6] was ruled out by a later ^{2}H NMR study via incorporation of deuterated mevalonates into ovalicin.[16] Furthermore, the proposed intermediate **10** has been synthesized,[14] but an attempt to incorporate the isotope-labeled **10** to ovalicin by feeding to *P. ovalis* was unsuccessful.[17] Thus, the biosynthesis of **3** may proceed via an alternative pathway, such as

dx.doi.org/10.1021/ja312503y | J. Am. Chem. Soc. 2013, 135, 4616−4619

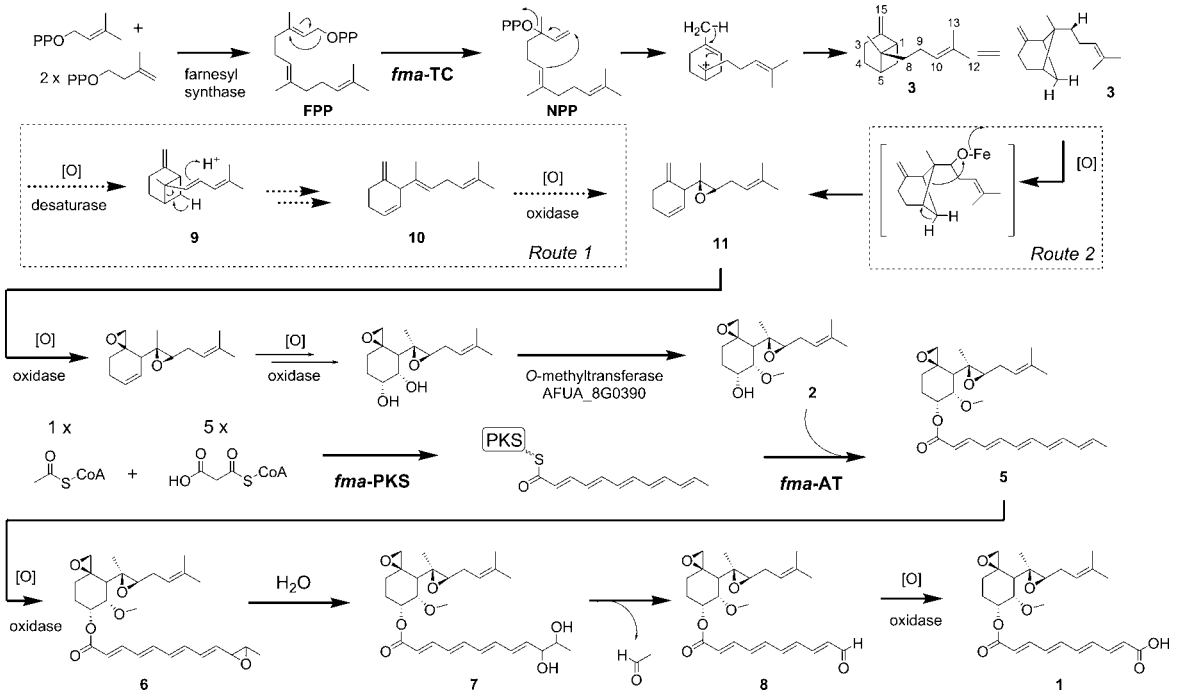

Figure 4. Putative biosynthetic pathway of **1**.

that in route 2, in which an iron-dependent oxygenase, such as AFUA_8G00480, can hydroxylate **3** at C8, followed by rearrangement of the oxo-Fe intermediate to yield the epoxide **11** directly. A second epoxidation step, dihydroxylation, and *O*-methylation afford **2**. The conversion of **2** to exclusively **5** in the presence of *fma*-PKS and -AT, instead of the shorter decatetraenoate ester, is surprising. This suggests that formation of the dioic acid moiety in **1** may proceed via an oxidative cleavage of the terminal alkene in **5** to yield the dioic acid in **1** (Figure 4). Oxidative cleavage of alkene by a single oxygenase (either via a mono- or dioxygenase mechanism) is well-documented, as exemplified by the carotenoid cleavage oxygenases.[24] We propose the terminal alkene in **5** can undergo oxidative C−C bond cleavage through epoxidation to yield **6**, then hydrolysis to afford a vicinal diol **7**, which can undergo cleavage to yield the aldehyde **8** and acetaldehyde. Final oxidation of **8** then yields **1**.

Identification of the *fma* gene cluster enabled us to take a deeper look into the nearby genes for a possible terpene cyclase that may synthesize **3**. The locus AFUA_8G00520 (Af0520) was originally annotated as encoding an integral membrane protein with 173 amino acids. Further conserved domain analysis and re-annotation revealed that the region encodes a 267-residue protein that contains a PFAM01040 UbiA prenyltransferase (*p*-hydroxybenzoic acid oligoprenyltransferase) domain (SI). However, Af0520 exhibits only low homology (14−19% protein identity/~30% similarity) to the *E. coli* UbiA, *S. cerevisiae* COQ2 (UbiA homologue), and other known fungal membrane-bound prenyltransferases,[25] and has almost no homology to the recently discovered membrane-bound cyclase involved in biosynthesis of pyripyropene A[19] and 3,5-dimethylorsellinic acid (DMOA)-derived meroterpenoids.[26] Protein structure prediction indicates that Af520 is comprised of six transmembrane helices (Figure S5). Although they share only minimal homology, the *E. coli* UbiA prenyltransferase has been shown to be structurally related to Class I terpene synthases, both of which belong to all-α-helical proteins, and a 3D model for UbiA has been built based on *N. tabacum epi*-aristolochene synthases.[27] Indeed, homology to the α-helical structures of FPP synthase can also be detected for Af520 (Figure S5). We therefore reasoned that Af520 may encode a new type of terpene cyclases (named *fma*-TC) that may be responsible for the cyclization of FPP to **3**.

To investigate the involvement of *fma*-TC in the biosynthesis of **1**, we inactivated the gene through homologous recombination in *A. fumigatus aku*^{BKU80} strain. Indeed, the resulting ΔAf520 mutant lost the abilities to produce **1** compared to the isogenic control strain (Figure 5A). When **3** or **5** was supplied to the ΔAf520 strain at 100 or 40 mg/L, respectively, production of **1** was restored (Figure 5A), thereby directly linking *fma*-TC to the generation of **3** in *A. fumigatus*. Next, we cloned the cDNA of *fma*-TC (for construction of the correct cDNA, see SI) into a yeast 2μ expression vector and transformed it into BJ5464-NpgA. We reasoned that the recombinant *fma*-TC may be able to utilize the endogenous yeast FPP to produce **3**. Gratifyingly, after 4 days of culturing followed by extraction with hexane, we were able to detect the formation of a highly hydrophobic compound that had the same mass (204.2) as **3** using GC-MS (Figure 5B) from the yeast strain. In contrast, a control strain that does not express *fma*-TC did not produce any detectable amount of **3**. Larger scale culturing enabled us to isolate the compound at a purified titer of 46 mg/L. GC-MS fragmentation (Figure 5E) and NMR analyses (SI) confirmed that the yeast-derived compound is identical to **3**.[11,14] Lastly, we prepared microsomal fractions from the yeast strain that overexpressed *fma*-TC. When provided with 1 mM FPP and 5 mM $MgCl_2$, we were able to detect nearly sole formation of **3** in vitro. With ~250 μg/mL microsomal protein, the apparent rate of formation of **3** was estimated to be ~4 μM/min. Like UbiA,[27,28] the *fma*-TC-containing microsomes were unable to convert FPP into **3** in the absence of Mg^{2+}. Collectively, these results confirmed that the membrane-bound *fma*-TC is indeed the terpene cyclase that produces **3** as an intermediate for **1** in *A. fumigatus*.

Based on the type of catalytic reaction and protein fold, terpene cyclases can be generally categorized in the Class I

dx.doi.org/10.1021/ja312503y | J. Am. Chem. Soc. 2013, 135, 4616−4619

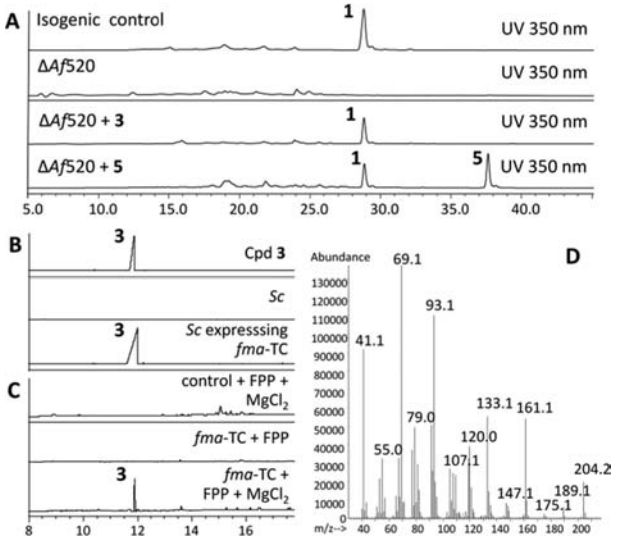

Figure 5. Verification of the function of Af520 encoding *fma*-TC and the synthesis of **3**. (A) HPLC analysis of metabolites produced by *A. fumigatus* upon deletion of Af520, and chemical complementation of ΔAf520 with **3** or **5**. (B) GC-MS detection of **3** upon expression of *fma*-TC in *S. cerevisiae* BJ5464-NpgA. *Sc* control is untransformed BJ5464-NpgA. (C) GC-MS analysis of in vitro assays with FPP and yeast microsomes containing *fma*-TC. (D) GC-MS spectrum of **3** from yeast.

(ionization of the allylic diphosphate ester to generate an allylic cation) or Class II (protonation-initiated) terpene cyclases.[29] The recently characterized cyclases in pyripyropene[19] and in DMOA-derived meroterpenoid pathways[26] formed a new family of membrane-bound terpene cyclases that catalyze a Class II-type cyclization. On the other hand, *fma*-TC (*β-trans*-bergamotene synthase) uncovered in this study is a new membrane-bound terpene cyclase that catalyzes a Class I-type reaction. Interestingly, plants are known to produce several bergamotenes using soluble Class I terpene cyclases.[30] Thus, this study demonstrates another intriguing case of convergent evolution of seemingly unrelated enzymes from different kingdoms to synthesize similar products.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental details and spectroscopic data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

yitang@ucla.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

H.-C.L. is supported by National Science Council of Taiwan (102-2917-I-564-008). This work was supported by NIH 1R01GM085128 and 1DP1GM106413 to Y.T., and by Northern Illinois University to A.M.C. NMR instrumentation was supported by the NSF equipment grant CHE-1048804. We deeply thank Prof. David Cane for enlightening discussions on the biosynthetic route. We thank Prof. John Porco Jr. and Dr. Alexander Grenning of CMLD-BU for the generous gifts of **1** and **2**. We also thank Dr. Kangjian Qiao for helpful discussions.

■ REFERENCES

(1) McCowen, M. C.; Callender, M. E.; Lawlis, J. F. *Science* **1951**, *113*, 202.
(2) Killough, J. H.; Magill, G. B.; Smith, R. C. *Science* **1952**, *115*, 71.
(3) Molina, J. M.; Tourneur, M.; Sarfati, C.; Chevret, S.; de Gouvello, A.; Gobert, J. G.; Balkan, S.; Derouin, F. *N. Engl. J. Med.* **2002**, *346*, 1963.
(4) Sin, N.; Meng, L. H.; Wang, M. Q. W.; Wen, J. J.; Bornmann, W. G.; Crews, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6099.
(5) Lefkove, B.; Govindarajan, B.; Arbiser, J. L. *Exp. Rev. Anti Infect. Ther* **2007**, *5*, 573.
(6) Birch, A. J.; Hussain, S. F. *J. Chem. Soc., Perkin Trans. 1* **1969**, *11*, 1473.
(7) (a) Yamaguchi, J.; Hayashi, Y. *Chem. Eur. J.* **2010**, *16*, 3884. (b) Corey, E. J.; Snider, B. B. *J. Am. Chem. Soc.* **1972**, *94*, 2549. (c) Yamaguchi, J.; Toyoshima, M.; Shoji, M.; Kakeya, H.; Osada, H.; Hayashi, Y. *Angew. Chem., Int. Ed.* **2006**, *45*, 789.
(8) Balthaser, B. R.; Maloney, M. C.; Beeler, A. B.; Porco, J. A., Jr.; Snyder, J. K. *Nat. Chem.* **2011**, *3*, 969.
(9) (a) Kruger, E. A.; Figg, W. D. *Exp. Opin. Invest. Drugs* **2000**, *9*, 1383. (b) Yin, S. Q.; Wang, J. J.; Zhang, C. M.; Liu, Z. P. *Curr. Med. Chem.* **2012**, *19*, 1021.
(10) Cane, D. E.; Mcilwaine, D. B. *Tetrahedron Lett.* **1987**, *28*, 6545.
(11) Nozoe, S.; Kobayashi, H.; Morisaki, N. *Tetrahedron Lett.* **1976**, 4625.
(12) Cane, D. E.; Mcilwaine, D. B.; Oliver, J. S. *J. Am. Chem. Soc.* **1990**, *112*, 1285.
(13) Cane, D. E.; Mcilwaine, D. B.; Harrison, P. H. M. *J. Am. Chem. Soc.* **1989**, *111*, 1152.
(14) Cane, D. E.; King, G. G. S. *Tetrahedron Lett.* **1976**, 4737.
(15) Cane, D. E.; Levin, R. H. *J. Am. Chem. Soc.* **1976**, *98*, 1183.
(16) Cane, D. E.; Buchwald, S. L. *J. Am. Chem. Soc.* **1977**, *99*, 6132.
(17) Cane, D. In *Biosynthesis of isoprenoid compounds*; Porter, J., Spurgeon, S., Eds.; Wiley: New York, 1981; Vol. *1*, p 283.
(18) Nierman, W. C.; Pain, A.; Anderson, M. J.; Wortman, J. R.; Kim, H. S.; Arroyo, J.; Berriman, M.; Abe, K.; Archer, D. B.; Bermejo, C. *Nature* **2005**, *438*, 1151.
(19) Itoh, T.; Tokunaga, K.; Matsuda, Y.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. *Nat. Chem.* **2010**, *2*, 858.
(20) Da Silva Ferreira, M. E.; Kress, M. R.; Savoldi, M.; Goldman, M. H.; Hartl, A.; Heinekamp, T.; Brakhage, A. A.; Goldman, G. H. *Eukaryot. Cell* **2006**, *5*, 207.
(21) Ma, S. M.; Li, J. W.; Choi, J. W.; Zhou, H.; Lee, K. K.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C. *Science* **2009**, *326*, 589.
(22) Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J. Am. Chem. Soc.* **2009**, *131*, 8388.
(23) Winter, J. M.; Sato, M.; Sugimoto, S.; Chiou, G.; Garg, N. K.; Tang, Y.; Watanabe, K. *J. Am. Chem. Soc.* **2012**, *134*, 17900.
(24) (a) Leuenberger, M. G.; Engeloch-Jarret, C.; Woggon, W. D. *Angew. Chem., Int. Ed.* **2001**, *40*, 2614. (b) Schmidt, H.; Kurtzer, R.; Eisenreich, W.; Schwab, W. *J. Biol. Chem.* **2006**, *281*, 9845.
(25) (a) Itoh, T.; Tokunaga, K.; Radhakrishnan, E. K.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. *ChemBioChem* **2012**, *13*, 1132. (b) Lo, H. C.; Entwistle, R.; Guo, C. J.; Ahuja, M.; Szewczyk, E.; Hung, J. H.; Chiang, Y. M.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2012**, *134*, 4709.
(26) Matsuda, Y.; Awakawa, T.; Itoh, T.; Wakimoto, T.; Kushiro, T.; Fujii, I.; Ebizuka, Y.; Abe, I. *ChemBioChem* **2012**, *13*, 1738.
(27) Brauer, L.; Brandt, W.; Schulze, D.; Zakharova, S.; Wessjohann, L. *ChemBioChem* **2008**, *9*, 982.
(28) Melzer, M.; Heide, L. *Biochim. Biophys. Acta* **1994**, *1212*, 93.
(29) Gao, Y.; Honzatko, R. B.; Peters, R. J. *Nat. Prod. Rep.* **2012**, *29*, 1153.
(30) (a) Sallaud, C.; Rontein, D.; Onillon, S.; Jabes, F.; Duffe, P.; Giacalone, C.; Thoraval, S.; Escoffier, C.; Herbette, G.; Leonhardt, N. *Plant Cell* **2009**, *21*, 301. (b) Landmann, C.; Fink, B.; Festner, M.; Dregus, M.; Engel, K. H.; Schwab, W. *Arch. Biochem. Biophys.* **2007**, *465*, 417.