JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Synthesis of (*R*)-Mellein by a Partially Reducing Iterative Polyketide Synthase

Huihua Sun,[†] Chun Loong Ho,[†] Feiqing Ding,[‡] Ishin Soehano,[†] Xue-Wei Liu,[‡] and Zhao-Xun Liang*[,†]

[†]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551
[‡]School of Mathematics and Physics, Nanyang Technological University, 21 Nanyang Link, Singapore 637371

Supporting Information

ABSTRACT: Mellein and the related 3,4-dihydroisocoumarins are a family of natural products with interesting biological properties. The mechanisms of dihydroisocoumarin biosynthesis remain largely speculative today. Here we report the synthesis of mellein by a partially reducing iterative polyketide synthase (PR-PKS) as a pentaketide product. Remarkably, despite the head-to-tail homology shared with several fungal and bacterial PR-PKSs, the mellein synthase exhibits a distinct keto reduction pattern in the synthesis of the pentaketide. We present evidence to show that the ketoreductase (KR) domain alone is able to recognize and differentiate the polyketide intermediates, which provides a mechanistic explanation for the programmed keto reduction in these PR-PKSs.

Mellein (3,4-dihydro-8-hydroxy-3-methylisocoumarin, **1**) and other dihydroisocoumarins are widespread secondary metabolites isolated from various organisms. With their notable biological activities, some of the dihydroisocoumarins are being considered as drug leads (see Figure 1).[1] Mellein, isolated from fungi and insect secretions (as a trail pheromone) and first discovered in 1933, exhibits fungicidal, antibacterial, and HCV protease-inhibitory properties.[2,3] Ochratoxin A (**3**) is a well-known fungal toxin that contains a chlorinated dihydroisocoumarin.[4,5] An immunomodulatory effect of ochratoxin A on a human macrophage cell line has been established. Bacilosarcin A (**7**) originating from a marine octocoral-associated bacterium exhibits potent antiplasmodial activity.[6,7] Hydrangenol (**8**), phyllodulcin (**9**) and thunberginol A (**10**) are structurally related dihydroisocoumarins isolated from the leaves of *Hydrangea macrophylla* var. *thunbergii* that promote adipogenesis of 3T3-L1 cells and exhibit antidiabetic properties.[8,9] AI-77-B (**11**) isolated from the culture broth of *Bacillus pumilus* AI-77 is endowed with gastroprotective and antiulcerogenic properties without exhibiting adverse effects on the central nervous system.[10,11] However, despite the fact that some of the naturally occurring dihydroisocoumarins have been known for many decades, the enzymes responsible for the biosynthesis of these compounds remain to be identified. It has been proposed that 6-methoxymellein (**4**) and ochratoxin A are synthesized by polyketide synthases (PKSs), but the sequence and domain composition of the PKSs have not been fully established.[12−15] The only system characterized to date is the type-I modular PKS responsible for the synthesis of the 7-methylated dihydroisocoumarin ring of ajudazols.[16]

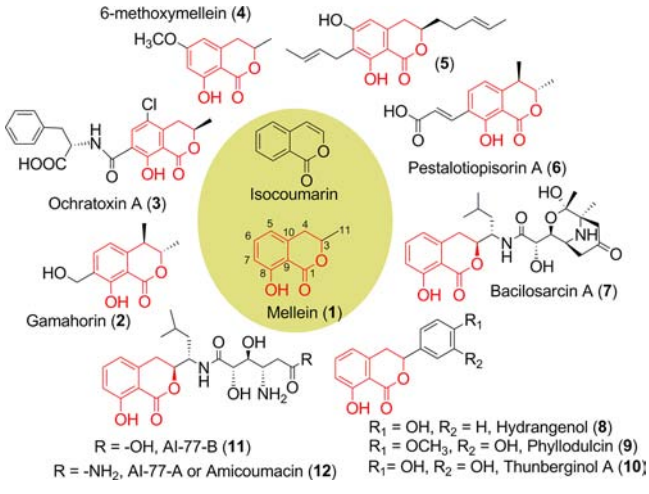

Figure 1. Isocoumarin, mellein, and representatives of naturally occurring 3,4-dihydroisocoumarins.

Saccharopolyspora erythraea is a mycelium-forming actinomycete that produces the macrolide antibiotic erythromycin A. The modular PKS responsible for the biosynthesis of erythromycin A has been a model system for studying polyketide synthesis. Sequencing of the genome of *S. erythraea* NRRL23338 revealed a dozen additional PKS genes.[17] The products of many of these PKSs remain unidentified despite extensive fermentation experiments. One of the uncharacterized orphan PKS genes is *SACE5532* (or *pks8*), which encodes a single-module PKS that shares head-to-tail sequence homology with several fungal and bacterial type-I partially reducing iterative PKSs (PR-PKSs) for aromatic polyketide biosynthesis (Figure 2a and Figure S1 in the Supporting Information). Among these PR-PKSs, ChlB1, MdpB, and PokM1 synthesize the 6-methylsalicyclic acid (6-MSA) moieties of chlorothricin, maduropeptin and pactamycin, respectively;[18−20] NcsB synthesizes the 2-hydroxy-5-methyl-NPA (NPA = 1-naphthoic acid) moiety of neocarzinostatin;[21] and AziB synthesizes the 5-methyl-NPA moiety of azinomycin B.[22] Accordingly, SACE5532 is predicted to contain a ketosynthase (KS), an acyltransferase (AT), a thioester hydrolase (TH),[23] a ketoreductase (KR) domain, and an acyl carrier protein (ACP) domain.

Received: May 21, 2012
Published: July 13, 2012

dx.doi.org/10.1021/ja304905e | *J. Am. Chem. Soc.* 2012, 134, 11924−11927

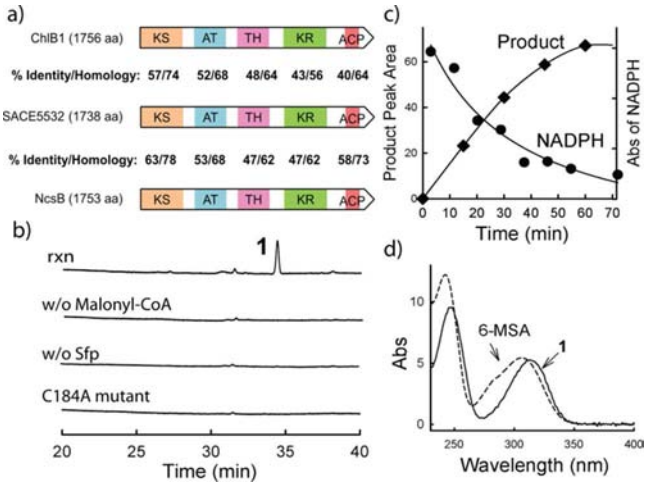

Figure 2. Domain organization and enzymatic activity of SACE5532. (a) Domain organization of SACE5532 and the sequence identity/homology of the domains shared with ChlB1 and NcsB. b) In vitro enzymatic activity of SACE5532 analyzed by HPLC. (c) Time-dependent depletion of NADPH and formation of enzymatic product in the enzymatic assay. (d) Absorption spectra of 6-MSA and the enzymatic product of SACE5532.

To identify the product of SACE5532, the recombinant SACE5532 was cloned and coexpressed with the phosphopantetheinyl transferase (PPTase) Sfp in *Escherichia coli* to generate the phosphopantetheinylated protein. Coexpression with Sfp not only increases the soluble fraction of the recombinant protein but also prevents the formation of high-molecular-weight oligomers (data not shown). The enzymatic activity of the recombinant SACE5532 was examined by in vitro assay. When SACE5532 was added to the reaction buffer containing acetyl-CoA, malonyl-CoA, and NADPH, the absorbance of NADPH at 340 nm began to decline immediately. We observed a faster depletion of NADPH when the PKS was preincubated with Sfp and coenzyme A, indicating that the protein is only partially phosphopantetheinylated by Sfp in *E. coli* cells. Product analysis by HPLC revealed that a single new species, **1**, was produced by the enzymatic reaction (Figure 2b and Figure S3). When acetyl-CoA or malonyl-CoA was omitted from the reaction mixture, the new species was not detected. No product was observed when SACE5532 was either not coexpressed or not pretreated with Sfp. When the putative catalytic residue Cys^{184} in the KS domain was replaced by an alanine residue, no product was detected either. A time-dependent enzymatic assay showed that the amount of product increased steadily with the consumption of NADPH before reaching the plateau (Figure 2c). These experiments establish that the recombinant SACE5532 is catalytically active and produces a single enzymatic product. The absorption spectrum of the new species indicates that the species is neither 6-MSA nor an NPA derivative (Figure 2d), suggesting that SACE5532 produces a novel product.

To find out whether SACE5532 generates the same product in a cellular environment, the medium of the *E. coli* cells coexpressing SACE5532 and Sfp was extracted with organic solvent. HPLC analysis of the medium extract showed that **1** was also produced by the *E. coli* cells (Figure 3a). The LC−MS experiment revealed a mass of 178.0723 Da for **1** and suggested a molecular formula of $C_{10}H_{10}O_3$ (m/z 178.0630 calcd) (Figure 3b). Moreover, when [1,2,3-$^{13}C_3$]malonic acid was used in

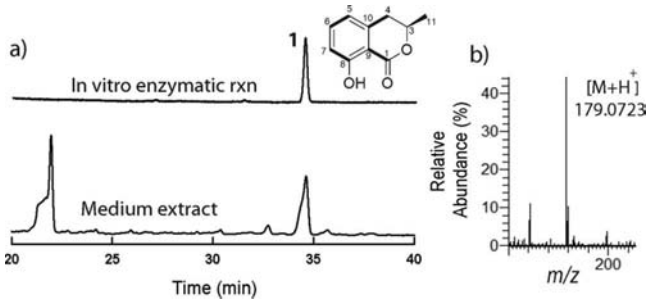

Figure 3. Production and characterization of the enzymatic product of SACE5532. (a) HPLC analysis of the formation of the enzymatic product by in vitro enzymatic reaction and in vivo coexpression of Sfp/SACE5532. The structure of the product, (*R*)-mellein [(*R*)-3,4-dihydro-8-hydroxy-3-methylisocoumarin] is shown. (b) Mass spectrum of the purified enzymatic product.

conjunction with the malonyl-CoA synthase MatB to produce ^{13}C-labeled malonyl-CoA in situ for the in vitro reaction,[24] the ^{13}C-labeled product was found to exhibit a molecular mass of 186.0992 Da (m/z 186.0780 calcd) (Figure S3). The difference of 8 Da between the labeled and non-labeled products suggests that the product is derived from a pentaketide intermediate with eight carbon atoms from malonic acid and the other two from the starter unit acetyl-CoA. A large-scale in vitro reaction was carried out, and the enzymatic product was purified by chromatography for NMR structure determination. Together with the mass spectrometry data, the NMR spectra established the product as mellein (see the NMR data in Figures S4−S6). The *R* configuration at C3 was established by comparing the optical specific rotation of the enzymatic product with that of the synthetic standard.[25] A survey of the literature suggested that (*R*)-(−)-mellein is the most common stereoisomer isolated from natural sources.

On the basis of the structure of the product and the chemistry of polyketide synthesis, we propose a biosynthetic mechanism for the novel mellein synthase (Figure 4). With the phosphopantetheinylated SACE5532, the AT domain loads the acyl groups from the acyl substrates onto the KS and ACP domains. A decarboxylative Claisen condensation catalyzed by the KS domain follows, generating the diketide intermediate.

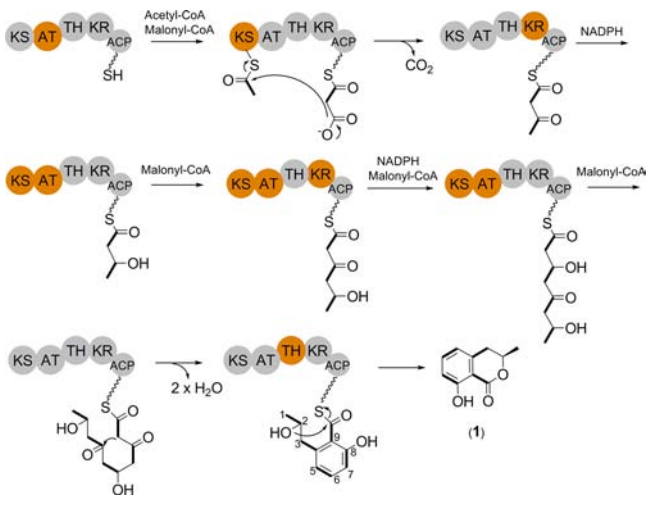

Figure 4. Proposed biosynthetic mechanism for the mellein-synthesizing SACE5532.

dx.doi.org/10.1021/ja304905e | J. Am. Chem. Soc. 2012, 134, 11924−11927

The loading of the malonyl group and condensation are repeated four times to generate a pentaketide intermediate covalently tethered to the ACP domain. During this process, the KR domain selectively reduces the keto group in the diketide and tetraketide intermediates but not the triketide and pentaketide intermediates. The pentaketide intermediate subsequently undergoes an aldol cyclization to furnish the aromatic structure through dehydration, and the TH domain catalyzes the release of (*R*)-mellein via a stereospecific cyclization process.

Although SACE5532 has the same domain composition and shares head-to-tail sequence homology with ChlB1, MdpB, PokM1, NcsB, and AziB, it generates a pentaketide product rather than the tetra- or hexaketide-derived 6-MSA and NPAs. More strikingly, the mechanism depicted in Figure 4 indicates that SACE5532 exhibits a completely different keto reduction pattern. As illustrated in Figure 5, while the other three PKSs

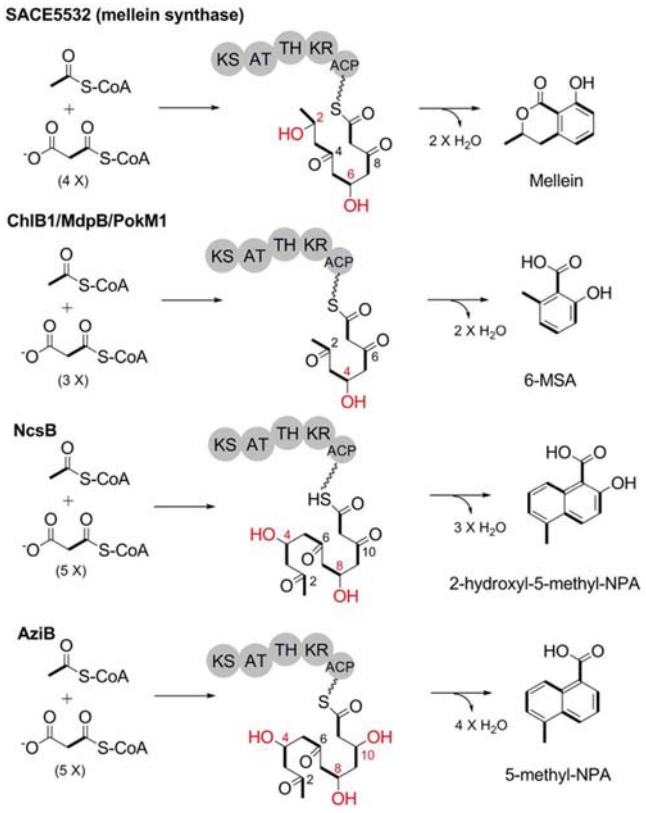

Figure 5. Keto reduction patterns for SACE5532 and the homologous PR-PKSs.

reduce the keto groups at C4, C8, and C10, SACE5532 reduces the keto groups at C2 and C6. In other words, while ChlB1, NcsB, and AziB reduce the *β*-keto group in the tri-, penta-, and hexaketide intermediates, SACE5532 reduces the keto group in the diketide and tetraketide intermediates. Supporting the proposed mechanism and diketide intermediates in mellein synthesis is the observation that SACE5532 is able to accept acetoacetyl-CoA or hydroxybutyryl-CoA as the starter unit to initiate the synthesis of mellein, albeit at lower rates (Figure 6a). Direct loading of the two diketide acyl groups presumably enables the enzyme to bypass the first condensation step.

According to the mechanisms illustrated in Figure 5, the keto group of the diketide intermediate is reduced by SACE5532 but not by ChlB1, NcsB, or AziB (also see Figure S8). How does

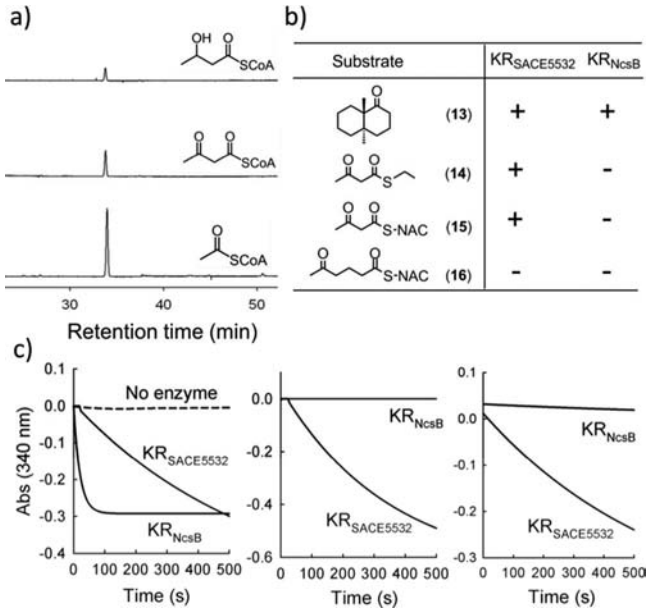

Figure 6. Substrate specificity of the KR domains of SACE5532 and NcsB. (a) HPLC analysis of the production of mellein by SACE5532 with different starter units. (b) Selective reduction of the *β*-keto group of the diketide analogues by the KR domain of SACE5532. (c) Ketoreductase activity of the KR domains of SACE5532 and NcsB toward (left) *trans*-1-decalone (**13**), (middle) acetoacetyl-SNAC (**15**), and (right) *S*-ethyl acetothioacetate (**14**).

the mellein synthase achieve such a completely different keto reduction pattern? Multiple sequence alignment suggests that the four KR domains contain all of the critical catalytic residues for a typical NADPH-binding KR domain (Figure S7). Secondary structure prediction and structural modeling also suggest that the KR domain of SACE5532 shares a similar overall structure with typical KR domains.[26,27] To investigate whether the KR domain is solely responsible for determining the keto reduction pattern, we cloned and expressed the standalone KR domains from SACE5532 (1170−1638 amino acids) and NcsB (1152−1641 amino acids) (Figure S2). Both KR domains are soluble and enzymatically active, as demonstrated by the reduction of *trans*-1-decalone (**13**), a nonspecific substrate used for assaying the activity of KR domains (Figure 6b,c).[28] Because the analogues of the tri-, tetra-, and pentaketide are chemically labile as a result of their tendency to form lactones,[29] we used two diketide analogues to test whether the two standalone KR domains exhibit any substrate preference. When the analogue of the diketide intermediate acetoacetyl-*N*-acetylcysteamine thioester (acetoacetyl-SNAC) (**15**) was used as the substrate, only $KR_{SACE5532}$ was capable of reducing the keto group, as evidenced by the depletion of NADPH (Figure 6c). No reduction was observed for KR_{NcsB} even after prolonged incubation. Notably, $KR_{SACE5532}$ also exhibits enzymatic activity toward one of the simplest diketide analogues, *S*-ethyl acetothioacetate (**14**), indicating that the KR domains are able to differentiate the diketide intermediates through the recognition of the acetoacetyl moiety. In another control experiment, no reduction was observed for 5-oxohexanoyl-SNAC (**16**) for either KR domain (Figure 6b and Figure S9), suggesting that the reduction of the *β*-keto group of the diketide by $KR_{SACE5532}$ is rather specific. Together, these results suggest that the programmed keto reduction in these PR-PKSs is achieved by

dx.doi.org/10.1021/ja304905e | *J. Am. Chem. Soc.* 2012, 134, 11924−11927

discrimination of the polyketide intermediates by the KR domain alone, that is, the KR domain of SACE5532 binds only the diketide and tetraketide intermediates in a productive configuration to allow hydride transfer from NADPH.

In summary, the experimental results presented here establish the polyketide origin of mellein and may facilitate the identification of the biosynthetic gene clusters for other dihydroisocoumarins. Since isocoumarins are usually synthesized by the KR-domain-lacking nonreducing PKSs (NR-PKSs) (Figure S10), the results indicate that a distinct iterative PR-PKS system has evolved for the synthesis of the structurally related dihydroisocoumarins. It is also interesting to note that for one of the fungal species that can produce mellein and ochratoxins, a separate PKS seems to be responsible for assembling the ochratoxins.[30] It remains to be seen whether the ochratoxin-synthesizing PKS bears any similarity to the mellein synthase reported here. Comparative studies of the KR domains have yielded insight into the mechanism of the programmed keto reduction by the PR-PKSs. The substrate specificity exhibited by the standalone KR domains toward diketide analogues strongly suggests that the KR domain is able to differentiate and selectively reduce the polyketide intermediates in its substrate-binding pocket. In conjunction with the recent insights into the mechanism of NR-PKSs and the determination of the methylation pattern and stereospecific keto reduction in fungal PR-PKSs,[31−35] The results further raise the hope of reprogramming and redesigning the iterative PKSs in the near future.

■ ASSOCIATED CONTENT

Supporting Information

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

■ AUTHOR INFORMATION

Corresponding Author

zxliang@ntu.edu.sg

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by a Tier II ARC Grant to Z.X.L. from MOE, Singapore.

■ REFERENCES

(1) Butler, M. S. *Nat. Prod. Rep.* **2008**, *25*, 475.
(2) Holler, U.; Konig, G. M.; Wright, A. D. *J. Nat. Prod.* **1999**, *62*, 114.
(3) Krohn, K.; Bahramsari, R.; Flörke, U.; Ludewig, K.; Kliche-Spory, C.; Michel, A.; Aust, H.-J.; Draeger, S.; Schulz, B.; Antus, S. *Phytochemistry* **1997**, *45*, 313.
(4) Bayman, P.; Baker, J. L.; Doster, M. A.; Michailides, T. J.; Mahoney, N. E. *Appl. Environ. Microbiol.* **2002**, *68*, 2326.
(5) Bennett, J. W.; Klich, M. *Clin. Microbiol. Rev.* **2003**, *16*, 497.
(6) Azumi, M.; Ogawa, K.; Fujita, T.; Takeshita, M.; Yoshida, R.; Furumai, T.; Igarashi, Y. *Tetrahedron* **2008**, *64*, 6420.
(7) Li, Y. X.; Xu, Y.; Liu, L. L.; Han, Z.; Lai, P. Y.; Guo, X. R.; Zhang, X. X.; Lin, W. H.; Qian, P. Y. *Mar. Drugs* **2012**, *10*, 319.
(8) Matsuda, H.; Shimoda, H.; Yamahara, J.; Yoshikawa, M. *Biol. Pharm. Bull.* **1999**, *22*, 870.
(9) Yoshikawa, M.; Ueda, T.; Shimoda, H.; Murakami, T.; Yamahara, J.; Matsuda, H. *Chem. Pharm. Bull.* **1999**, *47*, 383.
(10) Itoh, J.; Shomura, T.; Omoto, S.; Miyado, S.; Yuda, Y.; Shibata, U.; Inouye, S. *Agric. Biol. Chem.* **1982**, *46*, 1255.
(11) Shimojima, Y.; Hayashi, H.; Ooka, T.; Shibukawa, M.; Iitaka, Y. *Tetrahedron* **1984**, *40*, 2519.
(12) Kurosaki, F.; Itoh, M.; Yamada, M.; Nishi, A. *FEBS Lett.* **1991**, *288*, 219.
(13) Kurosaki, F. *Phytochemistry* **1996**, *41*, 1023.
(14) O'Callaghan, J.; Caddick, M. X.; Dobson, A. D. W. *Microbiology* **2003**, *149*, 3485.
(15) Bacha, N.; Atoui, A.; Mathieu, F.; Liboz, T.; Lebrihi, A. *Fungal Genet. Biol.* **2009**, *46*, 77.
(16) Buntin, K.; Rachid, S.; Scharfe, M.; Blöcker, H.; Weissman, K. J.; Müller, R. *Angew. Chem., Int. Ed.* **2008**, *47*, 4595.
(17) Oliynyk, M.; Samborskyy, M.; Lester, J. B.; Mironenko, T.; Scott, N.; Dickens, S.; Haydock, S. F.; Leadlay, P. F. *Nat. Biotechnol.* **2007**, *25*, 447.
(18) Shao, L.; Qu, X.-D.; Jia, X.-Y.; Zhao, Q.-F.; Tian, Z.-H.; Wang, M.; Tang, G.-L.; Liu, W. *Biochem. Biophys. Res. Commun.* **2006**, *345*, 133.
(19) Van Lanen, S. G.; Oh, T. J.; Liu, W.; Wendt-Pienkowski, E.; Shen, B. *J. Am. Chem. Soc.* **2007**, *129*, 13082.
(20) Lu, W. L.; Roongsawang, N.; Mahmud, T. *Chem. Biol.* **2011**, *18*, 425.
(21) Sthapit, B.; Oh, T. J.; Lamichhane, R.; Liou, K.; Lee, H. C.; Kim, C. G.; Sohng, J. K. *FEBS Lett.* **2004**, *566*, 201.
(22) Zhao, Q. F.; He, Q. L.; Ding, W.; Tang, M. C.; Kang, Q. J.; Yu, Y.; Deng, W.; Zhang, Q.; Fang, J.; Tang, G. L.; Liu, W. *Chem. Biol.* **2008**, *15*, 693.
(23) Moriguchi, T.; Kezuka, Y.; Nonaka, T.; Ebizuka, Y.; Fujii, I. *J. Biol. Chem.* **2010**, *285*, 15637.
(24) An, J. H.; Kim, Y. S. *Eur. J. Biochem.* **1998**, *257*, 395.
(25) Islam, M. S.; Ishigami, K.; Watanabe, H. *Tetrahedron* **2007**, *63*, 1074.
(26) Keatinge-Clay, A. T.; Stroud, R. M. *Structure* **2006**, *14*, 737.
(27) Zheng, J. T.; Taylor, C. A.; Piasecki, S. K.; Keatinge-Clay, A. T. *Structure* **2010**, *18*, 913.
(28) Østergaard, L. H.; Kellenberger, L.; Cortés, J.; Roddis, M. P.; Deacon, M.; Staunton, J.; Leadlay, P. F. *Biochemistry* **2002**, *41*, 2719.
(29) Castonguay, R.; He, W. G.; Chen, A. Y.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2007**, *129*, 13758.
(30) Harris, J. P.; Mantle, P. G. *Phytochemistry* **2001**, *58*, 709.
(31) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243.
(32) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *J. Am. Chem. Soc.* **2011**, *133*, 16635.
(33) Zhou, H.; Gao, Z. Z.; Qiao, K. J.; Wang, J. J.; Vederas, J. C.; Tang, Y. *Nat. Chem. Biol.* **2012**, *8*, 331.
(34) Crawford, J. M.; Korman, T. P.; Labonte, J. W.; Vagstad, A. L.; Hill, E. A.; Kamari-Bidkorpeh, O.; Tsai, S.-C.; Townsend, C. A. *Nature* **2009**, *461*, 1139.
(35) Liu, T.; Chiang, Y. M.; Somoza, A. D.; Oakley, B. R.; Wang, C. C. C. *J. Am. Chem. Soc* **2011**, *133*, 13314.