JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Identification and Characterization of the Chaetoviridin and Chaetomugilin Gene Cluster in *Chaetomium globosum* Reveal Dual Functions of an Iterative Highly-Reducing Polyketide Synthase

Jaclyn M. Winter,[†] Michio Sato,[‡] Satoru Sugimoto,[‡] Grace Chiou,[§] Neil K. Garg,[§] Yi Tang,*[,†,§] and Kenji Watanabe*[,‡]

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

[‡]Department of Pharmaceutical Sciences, University of Shizuoka, Shizuoka 422-8526, Japan

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

Supporting Information

ABSTRACT: We report the identification and characterization of the *caz* biosynthetic cluster from *Chaetomium globosum* and the characterization of a highly reducing polyketide synthase (PKS) that acts in both a sequential and convergent manner with a nonreducing PKS to form the chaetomugilin and chaetoviridin azaphilones. Genetic inactivation studies verified the involvement of individual *caz* genes in the biosynthesis of the azaphilones. Through in vitro reconstitution, we demonstrated the in vitro synthesis of chaetoviridin A from the pyranoquinone intermediate cazisochromene using the highly reducing PKS and an acyltransferase.

Polyketides produced by fungi possess a diverse array of structural features and biological activities.[1,2] These compounds are assembled by iterative polyketide synthases (IPKSs), which are megasynthases containing single copies of catalytic domains that are programmed to function repeatedly in different combinations.[3] Additional programming complexity and structural diversity are generated when multiple IPKSs function collaboratively, as illustrated in the biosynthesis of the azaphilones asperfuranone (**1**) and azanigerone A (**2**) (Figure 1). This partnership between tandem IPKSs can occur in a sequential manner in which the product of a highly reducing PKS (HR-PKS) is transferred downstream to a nonreducing PKS (NR-PKS) for further chain elongation, as observed with **1**,[4] or in a convergent manner in which the HR-PKS and NR-PKS function independently of one another and their individual products are combined, as demonstrated with **2**.[5] Understanding the functions of individual IPKSs, as well as how collaborative IPKSs are recruited to function together, is therefore an important step toward decoding the relationship between fungal IPKSs and their products. In this work, we discovered a single HR-PKS from *Chaetomium globosum* that can partner in both sequential and convergent fashions with an NR-PKS to synthesize two structurally distinct fragments that are incorporated into the antifungal chaetoviridin A (**3**) and cytotoxic chaetomugilin A (**4**) (Figure 1). This finding uncovers a previously unknown mode of collaboration among IPKSs and further demonstrates how a limited number of IPKSs can be combined to generate highly complex scaffolds.

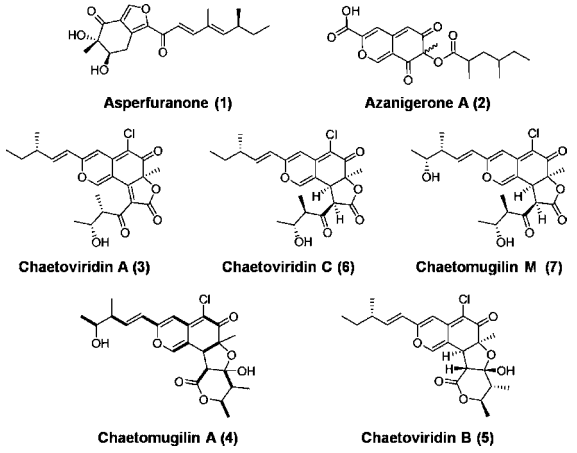

Figure 1. Structures of selected azaphilones and azaphilone-like molecules. The bold lines in chaetomugilin A (**4**) show enrichment from [1,2-$^{13}C_2$]acetate.

The common bicyclic pyranoquinone scaffold observed in the azaphilone family of polyketides is biosynthesized by an NR-PKS. The reduced substituents, such as those in **1** and **2**, are synthesized by a partnering HR-PKS found in each gene cluster.[4,5] *C. globosum* is a filamentous fungi that has been reported to produce a large variety of azaphilones, including chaetoviridin A (**3**),[6a] B (**5**),[6a,b] and C (**6**)[6a] as well as chaetomugilin A (**4**)[6c] and M (**7**)[6d] (Figure 1). Unlike **1** and **2**, the chlorinated compounds **3**, **6**, and **7** contain an angular lactone ring attached to the isochromenone core, whereas **4** and **5** contain a tetracyclic isochromenone−lactol−lactone structure.[6b,7] On the basis of their structural similarities, we propose that **3**−**7** are assembled via a common biosynthetic pathway. Feeding experiments with [1,2-$^{13}C_2$]acetate confirmed that the entire molecule of **4** is derived from a polyketide backbone (Figure 1 and Supplementary Figure 1 in the Supporting Information), which is consistent with previous isotopic experiments with structurally related azaphilones.[6a,8] Structural inspection of **3** and **4** revealed that in contrast to **1** and **2**, additional IPKS machinery must be required for their biosyn-

Received: September 12, 2012
Published: October 16, 2012

   dx.doi.org/10.1021/ja3090498 | *J. Am. Chem. Soc.* 2012, 134, 17900−17903

thesis. It can be speculated that assembly of the northern portions of **3** and **4** follows the biosynthetic logic of **1**, in which a sequential collaboration of an HR-PKS and an NR-PKS is required, whereas the assembly of the southern portions mimics that of **2**, in which the product of a different IPKS is combined with the northern portion in a convergent manner. Hence, we hypothesized that up to three IPKSs may be found in the biosynthetic gene cluster of **3**−**7**.

Using the HR-PKS and NR-PKS sequences from the gene clusters of **1** and **2** identified in *Aspergillus nidulans*[4] and *Aspergillus niger*,[5] respectively, we scanned the genome of *C. globosum* for putative azaphilone biosynthetic clusters containing two or more IPKSs. Three clusters containing an HR-PKS and an NR-PKS in close proximity were identified. Additionally, each cluster contained a predicted $FADH_2$-dependent halogenase that would introduce a chlorine atom onto the bicyclic core. Interestingly, none of these clusters contained three IPKSs as we hypothesized. Only one cluster (hereby named the *caz* cluster) encoded a putative acyltransferase (*caz*E) that would be required for transferring the convergent polyketide fragment (southern portion) to the pyranoquinone substrate (northern portion) to form **3**−**7**. This 65 kb *caz* cluster was annotated to encode 16 open reading frames, including genes encoding the HR-PKS (*caz*F) and NR-PKS (*caz*M) (Figure 2A and Supplementary Figure 2).

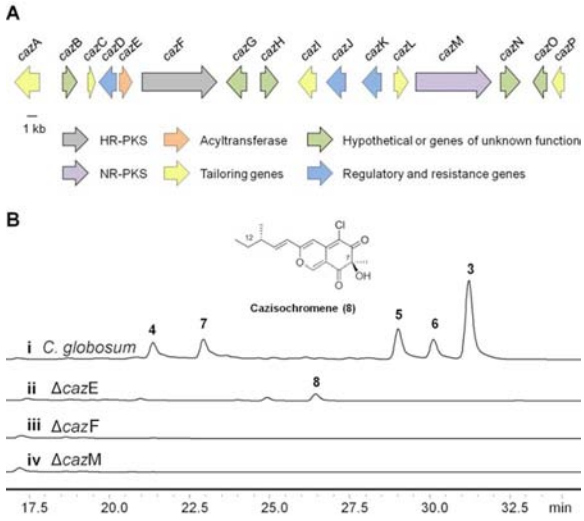

Figure 2. Genetic verification of the *caz* biosynthetic cluster in *C. globosum*. (A) Organization of the 65 kb *caz* cluster. (B) LC−MS analysis (360 nm) of (i) azaphilones produced by wild-type *C. globosum*, (ii) formation of **8** in the Δ*caz*E mutant, and (iii, iv) abolishment of the production of **3**−**7** in the (iii) Δ*caz*F and (iv) Δ*caz*M mutants.

To confirm the involvement of the *caz* cluster in the biosynthesis of **3**−**7**, gene inactivation of the NR-PKS *caz*M was accomplished via a fusion PCR gene-targeting cassette (Supplementary Figure 3).[9] The Δ*caz*M inactivation completely abolished production of the azaphilones. Inactivation of the HR-PKS *caz*F also completely abolished production of **3**−**7** without the recovery of any pyranoquinone intermediates, suggesting an essential role of both CazM and CazF in the synthesis of the northern portion of **3**−**7**. This is consistent with the collaborative mode of interaction between the two IPKSs, in which 4-methylhex-2-enoate **9** is the most likely product synthesized by CazF and is transferred to CazM via the starter-unit:ACP transacylase (SAT) domain of CazM as a starter unit for further chain elongation (see Figure 3). In this model, the NR-PKS is unable to produce any polyketide product in the absence of the starter unit.[10]

To obtain further confirmation of this mode of interaction between the two *caz* IPKSs, we inactivated the acyltransferase *caz*E that may be involved in connecting the northern and southern portions of **3**−**7**. We reasoned that inactivation of *caz*E should have no effect on the sequential collaboration between CazM and CazF. Indeed, the Δ*caz*E mutant was unable to produce **3**−**7** and instead accumulated an intermediate **8** with UV absorbance consistent with that of a pyranoquinone (Supplementary Figures 4−7). The structure of **8**, which we have named cazisochromene, was elucidated by 1D and 2D NMR spectroscopy (Figure 2B). The structure of the reduced substituent in **8** is consistent with that of **3**, **5**, and **6**, further confirming that **8** is an authentic intermediate in the *caz* pathway. However, the 12-hydroxyl derivative of **8** could not be detected in the culture, suggesting that the hydroxyl moiety present at this position in **4** and **7** may be installed by post-PKS oxygenases from the *caz* cluster rather than formed from incomplete reduction by the HR-PKS CazF. This hypothesis was supported by monitoring the time course of metabolite production and observing that the amount of **3** dramatically decreased while the amount of **4** increased over time (Supplementary Figure 8).

With **8** in hand, we propose the biosynthetic pathway of **3**−**7** shown in Figure 3. Following transfer of **9** from CazF to CazM, the polyketide undergoes four additional rounds of elongation and is methylated once, forming heptaketide intermediate **10**. While tethered to the acyl carrier protein (ACP), **10** undergoes a product template (PT)-domain-mediated regioselective cyclization via a C2−C7 aldol condensation followed by reductive release to form benzaldehyde **11**.[11,12] Chlorination of the aromatic ring catalyzed by the halogenase CazI and subsequent hydroxylation-catalyzed annulation at C7[5,13] by the predicted monooxygenase CazL then gives **8**. Since the *caz* cluster contains only two IPKSs instead of the initially hypothesized three, either CazF or CazM must provide the additional southern triketide fragment. We propose that CazF acts as a dual-functioning IPKS that supplies not only the reduced triketide **9** for CazM but also a more oxidized intermediate such as **12** for CazE (Figure 3). After CazE-catalyzed acylation, intramolecular aldol condensation of the newly formed β-keto ester yields **3.** Further modification of **3** through reduction (**3** → **6**), lactone cleavage and rearrangement, and C12 oxidation gives **4**−**7**.

To test the involvement of CazF and CazE in the conversion of **8** into **3**, both enzymes were cloned and expressed from heterologous hosts for in vitro reconstitution. CazE (54 kDa) was solubly expressed as an N-terminal octahistidine-tagged protein from *Escherichia coli* BL21(DE3)/pJW07637.[14] CazF was solubly expressed as a C-terminal hexahistidine-tagged protein from *Saccharomyces cerevisiae* BJ5464-NpgA/pJW07638.[15] The two proteins were purified to near homogeneity in yields of 52 and 9.5 mg/L, respectively (Supplementary Figures 9−11). The minimal PKS activity of CazF was evaluated by incubating the enzyme with 2 mM malonyl-CoA. Treatment of the reaction mixture with 1 M NaOH (base hydrolysis) followed by LC−MS analysis of the organic extract showed production of triacetic acid lactone **13** (Figure 4A), which is formed by spontaneous enolization and cyclization of the unreduced triketide.[10,16]

Intriguingly, in the absence of base hydrolysis, significantly lower levels of **13** were detected, indicating that the triketide may be fixed in an extended conformation by CazF to prevent

dx.doi.org/10.1021/ja3090498 | J. Am. Chem. Soc. 2012, 134, 17900−17903

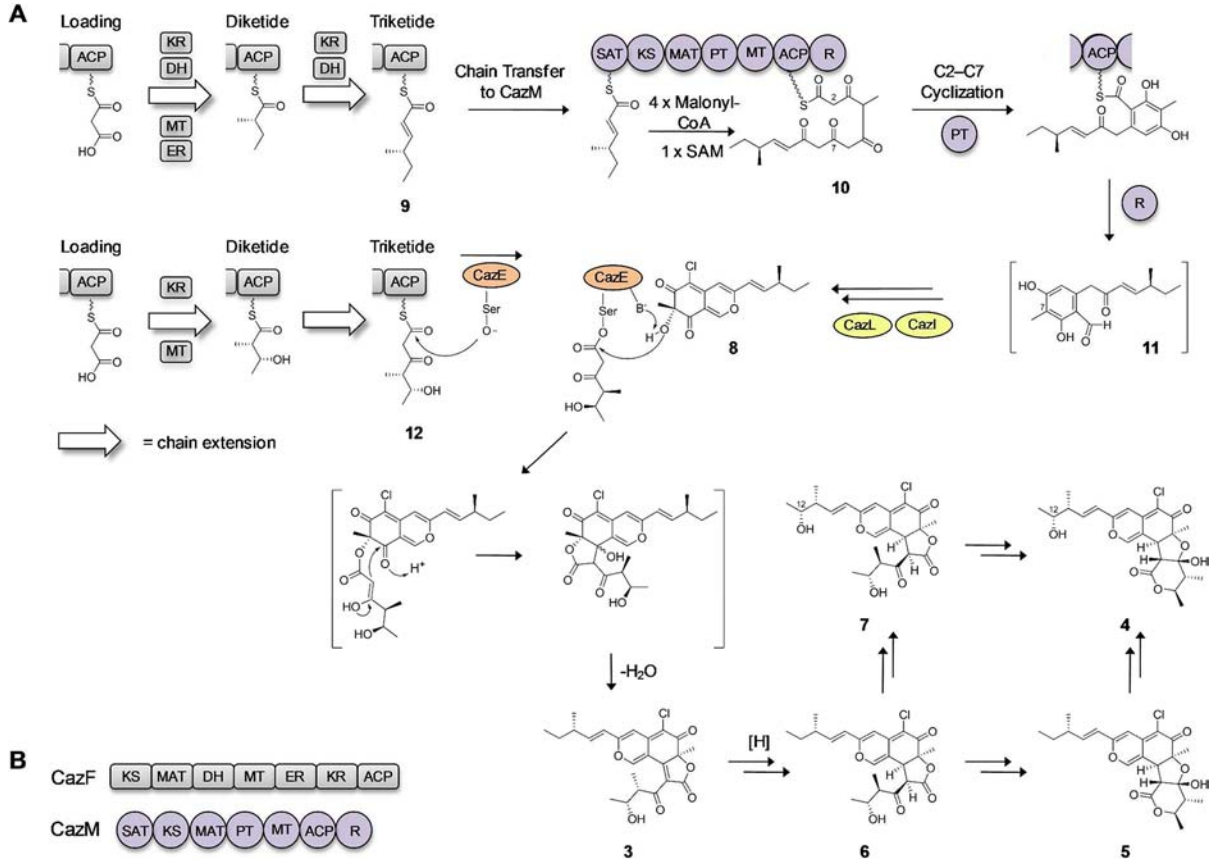

Figure 3. Proposed biosynthesis of the chaetoviridins and chaetomugilins from *C. globosum*. (A) Biosynthesis of **3**–**7** by the *caz* pathway. (B) Domain organization of the HR-PKS CazF and NR-PKS CazM. CazF consists of a ketosynthase (KS), malonyl-CoA:ACP acyltransferase (MAT), dehydratase (DH), methyltransferase (MT), enoyl reductase (ER), ketoreductase (KR), and acyl carrier protein (ACP). CazM consists of a starter-unit:ACP transacylase (SAT), KS, MAT, product template (PT), MT, ACP, and reductive domain (R).

lactonization and release. However, when CazE was added to the reaction, levels of **13** comparable to that of the base-hydrolyzed reaction were produced (Supplementary Figure 12). This suggested a likely interaction between CazE and CazF in which a conformational change in CazF after it interacts with CazE may expose the unreduced triketide to the aqueous environment, resulting in its subsequent lactonization and release.

We next examined the tailoring domains of CazF. The activity of the MT domain was assessed by adding 2 mM *S*-adenosyl-L-methionine (SAM) to CazF together with malonyl-CoA. Following incubation and base hydrolysis, dimethylated α-pyrone **14** (Figure 4A) was isolated and found to match a synthetically prepared standard (Supplementary Figures 13 and 14).[17] The formation of **14** is indicative of the correct timing of the MT domain following formation of diketide intermediates on CazF, which is required for the formation of both **9** and **12**. To assay the functions of the KR, DH, and ER domains, acetoacetyl-*S*-*N*-acetyl cysteamine (acetoacetyl-SNAC) was incubated with CazF and NADPH. LC–MS analysis of the organic extract showed accumulation of 3-hydroxybutanoyl-SNAC and butyryl-SNAC, confirming that all of the reductive domains are functional for the purified enzyme (Supplementary Figure 15).

After confirming that all of the domains of CazF are catalytically active, we attempted to synthesize **3** from **8**. Equimolar amounts of CazF and CazE were incubated with **8**, NADPH, SAM, and malonyl-CoA. LC–MS analysis of the organic extract revealed the formation of a new product whose retention time, UV profile, and isotopic mass ratio were identical to those of **3** (Figure 4B). When either CazF or CazE was omitted from the reaction, the formation of **3** could not be detected. Additionally, when only malonyl-CoA was added to the reaction or when NADPH was omitted, no new product formation was detected. However, when only SAM was omitted from the reaction, a new product with a shorter retention time than **3** and loss of a methyl group in mass was detected by the in vitro assay. On the basis of these observations, we assigned this product to be 4′-desmethylchaetoviridin A (**15**). Therefore, the in vitro reactions unequivocally confirmed the involvement of the HR-PKS in providing the southern portion of **3**–**7** and provided another example of an acyltransferase-mediated convergent synthesis of fungal polyketides.[18] The requirement of NADPH to produce the chaetoviridins suggests that CazE may possess substrate specificity toward a triketide reduced at the δ position, while the formation of **15** in the absence of SAM indicates that methylation at the γ position is not essential.

In summary, we have identified and characterized the *caz* biosynthetic cluster that is responsible for the production of the chaetoviridin and chaetomugilin azaphilones from *C. globosum*. Through in vivo gene inactivation experiments, we have established that the HR-PKS CazF most likely supplies a highly reduced methylated triketide product such as **9** to the NR-PKS CazM. Since CazM cannot be solubly expressed at this point, we cannot exclude the possibility that CazF transfers a polyketide product of different length to CazM as a starter unit. Additionally, through in vitro reconstitution experiments with the acyltransferase CazE, we established that CazF can also produce a more

dx.doi.org/10.1021/ja3090498 | *J. Am. Chem. Soc.* 2012, 134, 17900–17903

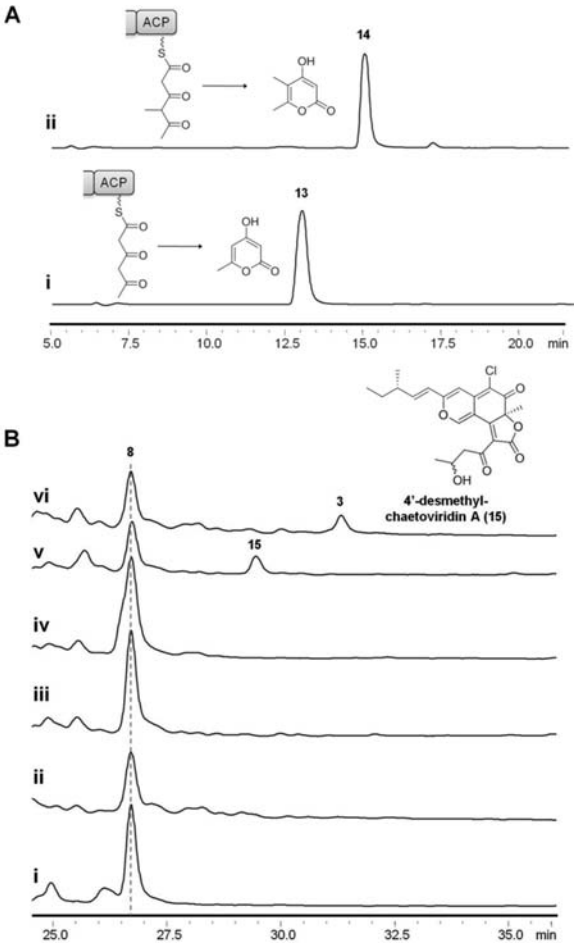

Figure 4. Reconstitution of CazE and CazF activity in vitro. (A) HPLC analysis (280 nm) of α-pyrones biosynthesized by CazF in the presence of (i) malonyl-CoA and (ii) malonyl-CoA and SAM. (B) HPLC analysis (360 nm) of chaetoviridins enzymatically synthesized in vitro when (i) CazF was incubated with **8**, malonyl-CoA, NADPH, and SAM; (ii) CazE was incubated with **8**, malonyl-CoA, NADPH, and SAM; (iii) CazE and CazF were incubated with malonyl-CoA; (iv) CazE and CazF were incubated with malonyl-CoA and SAM; (v) CazE and CazF were incubated with malonyl-CoA and NADPH; and (vi) CazE and CazF were incubated with malonyl-CoA, NADPH, and SAM.

oxidized triketide product, **12**, that is transacylated by CazE to a pyranoquinone intermediate, **8**, to afford **3**. Hence, CazF serves an unprecedented dual function in completing the biosynthesis of a polyketide product. The timing of substrate (**9** or **12**) transfer from CazF must be precisely orchestrated by the two interacting acyltrasferases in the form of the freestanding CazE and the SAT domain of CazM. Our work therefore unveils a new level of programming complexity among the fungal PKSs.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

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

■ AUTHOR INFORMATION

Corresponding Author

yitang@ucla.edu; kenji55@u-shizuoka-ken.ac.jp

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.M.W. was supported by a L'Oréal USA Fellowship for Women in Science. G.C. thanks the NSF for a graduate research fellowship (DGE-0707424). This work was supported by NIH (1R01GM085128 and 1DP1GM106413 to Y.T.) and by JSPS through the "Funding Program for Next Generation World-Leading Researchers" initiated by the Council for Science and Technology Policy (LS103), the Uehara Memorial Foundation, and the Mochida Memorial Foundation for Medical and Pharmaceutical Research (to K.W.). NMR instrumentation was supported by NSF equipment grant CHE-1048804.

■ REFERENCES

(1) Hoffmeister, D.; Keller, N. P. *Nat. Prod. Rep.* **2007**, *24*, 393.
(2) Brase, S.; Encinas, A.; Keck, J.; Nising, C. F. *Chem. Rev.* **2009**, *109*, 3903.
(3) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010.
(4) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965.
(5) Zabala, A.; Xu, W.; Chooi, Y. H.; Tang, Y. *Chem. Biol.* **2012**, *19*, 1049.
(6) (a) Takahashi, M.; Koyama, K.; Natori, S. *Chem. Pharm. Bull.* **1990**, *38*, 625. (b) Kingsland, S. R.; Barrow, R. A. *Aust. J. Chem.* **2009**, *62*, 269. (c) Muroga, Y.; Takeshi, Y.; Atsushi, N.; Reiko, T. *J. Antibiot.* **2008**, *61*, 615. (d) Muroga, Y.; Yamada, T.; Numata, A.; Tanaka, R. *Tetrahedron* **2009**, *65*, 7580.
(7) (a) Achard, M.; Beeler, A. B.; Porco, J. A., Jr. *ACS Comb. Sci.* **2012**, *14*, 236. (b) Germain, A. R.; Bruggemeyer, D. M.; Zhu, J.; Genet, C.; O'Brien, P.; Porco, J. A., Jr. *J. Org. Chem.* **2011**, *76*, 2577. (c) Musso, L.; Dallavalle, S.; Merlini, L.; Bava, A.; Nasini, G.; Penco, S.; Giannini, G.; Giommarelli, C.; De Cesare, A.; Zuco, V.; Vesci, L.; Pisano, C.; Castorina, M.; Milazzo, F.; Cervoni, M. L.; Dal Piaz, F.; De Tommasi, N.; Zunino, F. *Bioorg. Med. Chem.* **2010**, *18*, 6031.
(8) (a) Ogihara, J.; Kato, J.; Oishi, K.; Fujimoto, Y. *J. Biosci. Bioeng.* **2000**, *90*, 678. (b) Hajjaj, H.; Klaebe, A.; Loret, M. O.; Goma, G.; Blanc, P. J.; Francois, J. *Appl. Environ. Microbiol.* **1999**, *65*, 311. (c) Hajjaj, H.; Klaebe, A.; Goma, G.; Blanc, P. J.; Barbier, E.; Francois, J. *Appl. Environ. Microbiol.* **2000**, *66*, 1120.
(9) Szewczyk, E.; Nayak, T.; Oakley, C. E.; Edgerton, H.; Xiong, Y.; Taheri-Talesh, N.; Osmani, S. A.; Oakley, B. R. *Nat. Protoc.* **2006**, *1*, 3111.
(10) Zhou, H.; Qiao, K.; Gao, Z.; Meehan, M. J.; Li, J. W.; Zhao, X.; Dorrestein, P. C.; Vederas, J. C.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 4530.
(11) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243.
(12) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. *Chem. Commun.* **2007**, 4053.
(13) Davison, J.; al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 7642.
(14) Jez, J. M.; Ferrer, J. L.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Biochemistry* **2000**, *39*, 890.
(15) Lee, K. K.; Da Silva, N. A.; Kealey, J. T. *Anal. Biochem.* **2009**, *394*, 75.
(16) 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.; Tang, Y. *Science* **2009**, *326*, 589.
(17) Yokoe, H.; Mitsuhashi, C.; Matsuoka, Y.; Yoshimura, T.; Yoshida, M.; Shishido, K. *J. Am. Chem. Soc.* **2011**, *133*, 8854.
(18) Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J. Am. Chem. Soc.* **2009**, *131*, 8388.