P450-Mediated Coupling of Indole Fragments To Forge Communesin and Unnatural Isomers

Hsiao-Ching Lin,[†,§] Travis C. McMahon,[‡] Ashay Patel,[‡] Michael Corsello,[‡] Adam Simon,[‡] Wei Xu,[†] Muxun Zhao,[†] K. N. Houk,*[,‡] Neil K. Garg,*[,‡] and Yi Tang*[,†,‡]

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

[§]Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan

Ⓢ *Supporting Information*

ABSTRACT: Dimeric indole alkaloids are structurally diverse natural products that have attracted significant attention from the synthetic and biosynthetic communities. Here, we describe the characterization of a P450 monooxygenase CnsC from *Penicillium* that catalyzes the heterodimeric coupling between two different indole moieties, tryptamine and aurantioclavine, to construct vicinal quaternary stereocenters and yield the heptacyclic communesin scaffold. We show, via biochemical characterization, substrate analogues, and computational methods that CnsC catalyzes the C3-C3′ carbon-carbon bond formation and controls the regioselectivities of the pair of subsequent aminal bond formations to yield the communesin core. Use of ω-N-methyltryptamine and tryptophol in place of tryptamine led to the enzymatic synthesis of isocommunesin compounds, which have not been isolated to date.

Dimeric indole alkaloids are a large subset of plant and fungal natural products that have a wide range of biological activities. Examples include the tryptamine-derived calycanthaceous alkaloids such as (+)-calycanthine, epipolythiodiketopiperazines such as chaetocin, and the perophoramidine and communesin alkaloids.[1] A defining feature of these natural products is the presence of vicinal quaternary carbon stereocenters, which are particularly intriguing from both synthetic and biosynthetic perspectives.[2] Synthetic approaches to the vicinal quaternary stereocenters of these natural products have been developed and include Heck cyclization methodology, bis-(alkylation) chemistry, pericyclic reactions, and radical couplings.[2] The biosynthetic pathways to these natural products have been much less well-studied, and the enzymes responsible for vicinal quaternary stereocenter formation have not been revealed. Identifying the biocatalysts that form the C3-C3′ linkage is an important objective toward better understanding of indole dimer biosynthesis and how Nature can efficiently construct complex scaffolds.

We have specifically targeted the biosynthesis of the communesin alkaloids, given their remarkable structures and the vast interest in these compounds from the synthetic community.[3] The core structure of the communesins, as represented by the simplest isolated member, communesin K (3), consists of 7 interconnected rings, 2 aminal linkages, and 4

contiguous stereocenters. Other family members include communesins A (1) and B (2), which present additional synthetic challenges (Figures 1 and S1). Several total syntheses
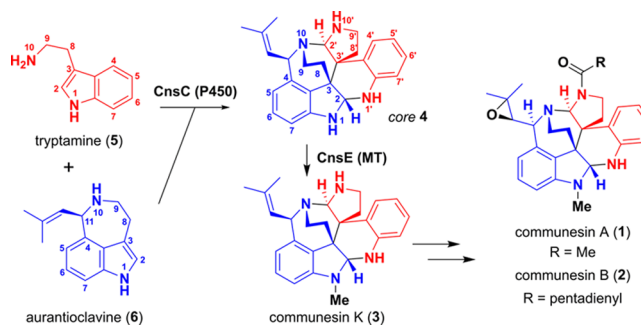

Figure 1. Formation of communesin core. A P450 (Pe-CnsC from *P. expansum*) has been implicated in the coupling of 5 and 6 into the core 4. Communesin K (3) is the simplest, stable communesin isolated, and is formed upon N-methylation of 4 by the methyltransferase Pe-CnsE.

have been reported.[4] With regard to communesin biosynthesis, Stoltz et al. proposed that the communesin core could be derived from the heterodimerization of tryptamine (5) and aurantioclavine (6), via direct C3-C3′ indole coupling followed by aminal bond formation or an exo inverse electron demand Diels-Alder reaction between 6 and the quinone methide imine derivative of 5.[2b] We recently identified the biosynthetic pathway of 1 and 2 in *Penicillium expansum* and verified that indeed, 5 and 6 are biosynthetic precursors to the core structure 4.[5] We found that a cytochrome P450 monooxygenase, CnsC, may be involved in the oxidative coupling between 5 and 6, as deletion of Pe-*cnsC* in *P. expansum* abolished communesin production and instead accumulated 6. Sequence analysis showed that CnsC displays low sequence homology to known P450s in the database (Figure S2). While CnsC must play a key role in the selective dimerization and generation of the core structure, numerous mechanistic questions remain to be answered: (1) are there other enzymes involved in the biosynthesis of 4, which requires the formation of three bonds (C3-C3′, C2-N1′, and C2′-N10) between 5 and 6?; (2) how are the sequences and regioselectivities of the bond-forming steps controlled to arrive

at 4?; and (3) what is the substrate specificity of CnsC toward two different indole building blocks in order to achieve the heterocoupling?

To investigate the activity of CnsC, the intron-free *cnsC* and the P450 redox partner cytochrome P450 oxidoreductase (Pe-CPR) from *P. expansum* were cloned under the ADH2 promoter and transformed into *Saccharomyces cerevisiae* strain BJ5464-NpgA.[6] However, Western blotting analysis of the membrane fraction showed very weak expression of the P450, and no new product could be detected when **5** and **6** were supplemented to the yeast culture. To find a CnsC alternative that may be better expressed in yeast, we sequenced the genome of *Penicillium rivulum*, which is also a reported producer of communesin.[3d] The *P. rivulum* communesin cluster is highly similar to that of the *cns* cluster in *P. expansum* (Table S2), including the P450 of interest, Pr-CnsC, which shows 84% sequence homology to Pe-CnsC and a different membrane anchoring region (Figure S3). Upon feeding **5** and **6** to the yeast strain expressing Pr-CnsC and Pe-CPR, we observed consumption of both substrates and production of trace amounts of **4** (m/z 385 $[M+H]^+$). The low signal of **4** is due to its instability from the two labile aminal linkages, which was also observed in knockout studies with *P. expansum*.[5] To further confirm formation of the core structure, we coexpressed the *N*-methyltransferase Pe-CnsE (Figure 1) with Pr-CnsC and Pe-CPR. Metabolite analysis of the yeast culture supplied with **5** and **6** showed clear accumulation of a new product that matches the retention time and m/z of the N1-methylated communesin K (**3**) (Figure S4A,ii). In contrast, untransformed yeast (Figure S4A,i) did not accumulate **3**.

Microsomal fractions containing the overexpressed Pr-CnsC and Pe-CPR were purified from 3-day cultures of the yeast expression strain. From the in vitro assay in which 0.5 mM **5**, 0.5 mM **6**, and 2 mM NADPH were incubated with microsomes (20 mg/mL total protein concentration), a clear product that corresponds to **4** emerged (Figure S4B,ii), albeit with low conversion. Introducing Pe-CnsE purified from *Escherichia coli* and 2 mM *S*-adenosyl-methionine (SAM) led to the formation of the expected **3**, fully consistent with the yeast biotransformation results (Figure S4B,iii). Collectively, these results demonstrate that the P450 CnsC alone is sufficient to oxidatively couple the two different indole-containing substrates together to yield the heptacyclic core of communesins. In vitro assays using either **5** or **6** alone did not lead to the formation of homodimeric products, or any other detectable products. CnsC is thus a rare example of a P450 that catalyzes C–C bond formation between two different substrates.[7] It is particularly intriguing that CnsC strategically unifies **5** and **6** with controlled formation of four new stereocenters and three new bonds. The most likely order of bond formation is the C3–C3′ coupling between **5** and **6**, followed by the two aminal bonds. It is also possible that the P450 active site orients the two substrates to promote aminal formation first, followed by the more difficult C3–C3′ bond formation.

To gain insight into the sequence of the bond forming steps and substrate specificities of CnsC, we tested enzyme activities toward different analogues of **5** and **6** using the in vitro reaction. When N1-methyl tryptamine or N1-methyl aurantioclavine was used in place of **5** or **6**, respectively, no product formation was detected, indicating the N1 hydrogens of both substrates are essential for the coupling reaction (Figure S7). In contrast, when N10-methyl tryptamine **7** was used, we observed the formation of two oxidatively coupled products **8** and **9** with the same mass as **3** (m/z 399 $[M+H]^+$) at a relative ratio of 5:1 (Figure 2A). To
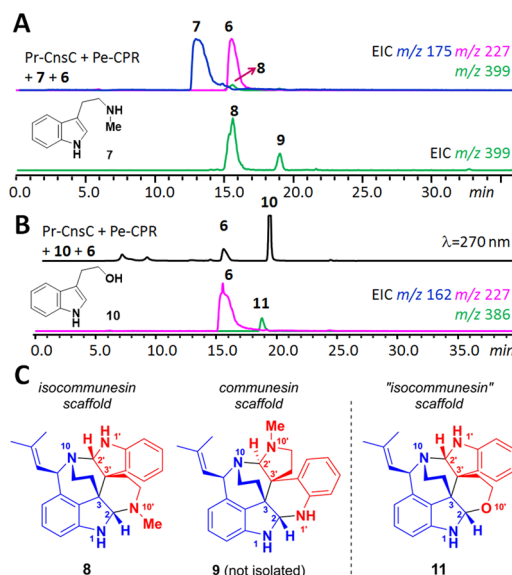

Figure 2. Formation of communesin and isocommunesin scaffolds in the presence of tryptamine analogues. (A) In vitro assay using **6** and **7** led to the synthesis of **8** and **9**; (B) assay using **6** and **10** led to the hemiaminal ether containing **11**; (C) structures of **8**, **9** (putative), and **11**.

elucidate the structures of **8** and **9**, a 4 L culture of the yeast strain coexpressing Pr-CnsC and Pe-CPR was grown for 3 days then concentrated 10-fold and supplemented with 5.0 mg each of **6** and **7**. Cultivation for 1 more day afforded sufficient quantities of **8**. The ^{1}H and ^{13}C NMR spectra of **8** showed characteristic signals of two aminals at δ 5.01 (H2′)/δ 84.8 (C2′) and δ 4.17 (H2)/δ 84.2 (C2) (Figures S8-S13; Table S3), and with other key HMBC correlations, the skeleton of **8** was established. The NOESY spectrum showed the correlations of H2′ to H8′β (δ 2.25), H8a (δ 2.65), and H9b (δ 3.42); H2 to H8′β and H8b (δ 2.08); H11 (δ 5.26) to H7′ (δ 6.67); and no correlation of H9 to H11 confirming the relative stereochemistry at the corresponding carbon atoms. On the basis of the *R* configuration of C11 in **6**,[8] the absolute stereochemistry of **8** was confirmed to be 2′*S*, 3′*S*, 3*R*, 2*S*, 11*R*. We were not able to isolate sufficient amount of **9** for full structural characterization. The shift in retention time and mass increase are indicative of the N10′ methylated form of **4**.

The structure of **8** is unexpected and represents an alternative way in which the two indole-containing building blocks **6** and **7** can be coupled. Compound **8** remains a heptacyclic alkaloid with the C3–C3′ linkage and four contiguous stereocenters, but with both indoline moieties remaining intact. In the communesin configuration, the two indole nitrogen atoms and the two ω-nitrogen atoms are each paired to form the two aminal linkages (N1–C2–N1′, N10–C2′–N10′). However, in **8** two different C–N bonds (N1–C2–N10′ and N1′–C2′–N10) are formed to give the bis(indoline) "isocommunesin" scaffold of **8**, which has not been isolated in any natural products to date. It is thus intriguing that N-methylation in **7** can drastically alter the products formed by CnsC.

To learn more about these substituent effects, we coupled other tryptamine analogues with **6** and monitored formation of adducts using LC/MS and selective ion monitoring. Surprisingly, CnsC displayed significant substrate tolerance toward analogues of **5**; product formation can be detected from substrates ranging in size between indole and ω-*N*-boc-tryptamine (Figure S5).
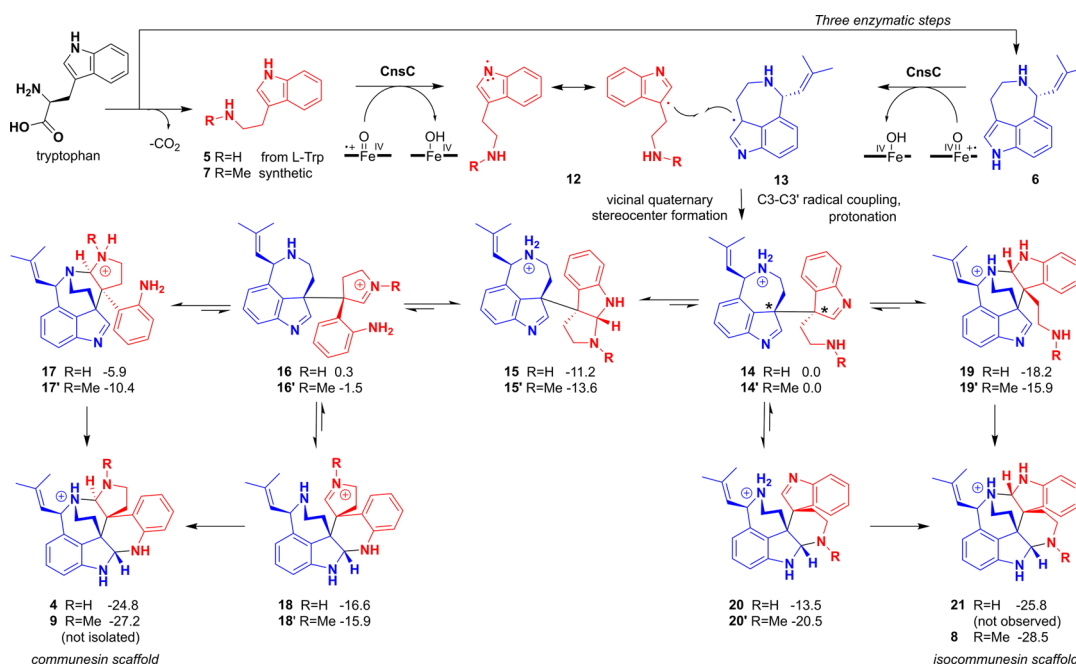
Figure 3. Proposed mechanism of CnsC catalyzed C3-C3′ coupling and formation of both communesin and isocommunesin scaffolds. **14** and **14**′ are the immediate products of the enzymatic oxidative coupling between tryptamine and aurantioclavine. Assuming N-protonation of substrate, intermediates, and products, the relative Gibbs free energies (in kcal mol^{-1}) of the aminal bond forming steps starting from **14** and **14**′ are shown.

Other C3-substituted indoles such as 3-methyl-indole, ω-N-dimethyl-tryptamine, and 3-(2-azidoethyl)-indole all yielded products with retention time suggestive of the isocommunesin structures (Figures S6). We were able to detect a single new product **11** (m/z 386 [M+H]$^{+}$) when tryptophol **10** was used as a substrate together with **6** (Figure 2B). Subsequent feeding of **6** and **10** into the yeast culture (6 mg each) led to the isolation of **11** (3 mg) for structural characterization (Figures S14-S18).

The ^{13}C NMR spectrum of **11** showed characteristic signals of one aminal at δ 83.0 (C2′), and one hemiaminal at δ 92.5 (C2). The HMBC correlations showed key correlations that confirmed the skeleton of **11** to be also of the isocommunesin scaffold. The absolute stereochemistries of the four contiguous stereocenters of **11** were confirmed to be analogous to **8** (2′S 3′S, 3R, 2S, 11R) based on NOESY data and the known R configuration of precursor **6** at C11. Interestingly, when only **6** was supplemented to the yeast culture expressing Pr-CnsC and Pe-CPR, **11** was also detected in the organic extracts. This is due to **10** being a known metabolite in *S. cerevisiae* that is involved in quorum sensing.[9] We also assayed the activities of CnsC toward N10-methyl analogues of **6**, but no coupled products were detected, suggesting the tolerance is low toward aurantioclavine analogues (Figure S7).

The reconstitution of CnsC activities and generation of either communesin or isocommunesin scaffolds provide significant insight into coupling between **5** and **6**. The common C3-C3′ bond in **4**, **8**, and **11** suggests that this coupling step takes place first in the active site of CnsC and staples the heterodimers together. Because C3-C3′ indole coupling is widely observed among alkaloid natural products and only a few mechanistic proposals exist, we computationally examined four possible mechanisms of coupling that can afford **4** (Figure S23). These mechanisms include radical addition, radical cation addition, electrophilic aromatic addition, and radical combination. The radical addition mechanism involves the intermediacy of a tryptaminyl radical formed by CnsC-catalyzed N-H hemolysis, followed by addition to **6**. Computations show that this

mechanism has a 10^5-fold intrinsic preference for the formation of a C3-C2′ adduct, which was not observed experimentally (Figure S24). Similarly, reaction of a tryptaminyl radical cation (generated by CnsC-catalyzed single electron transfer) with **6** is predicted to also lead to facile formation of a C3-C2′ product (Figure S26). C-C bond formation via the capture a tryptaminyl cation, the product of the two-electron oxidation of **5** by **6** is predicted to be nonselective (Figure S25). None of these mechanisms are deemed likely given the exclusive C3-C3′ coupling observed here. Further discussion and details of this computational work are provided in the Supporting Information.

In contrast, the combination of two indole C3 radicals **12** and **13** to form the vicinal quaternary stereocenters and yield the product **14** must exclusively form the C3-C3′ regioisomer (Figure 3). Both radicals can be generated by CnsC-catalyzed abstraction of the indole N1 hydrogen to give the aza-allyl radical that has highest spin density at C3. Such one-electron oxidation that initiates at the indole NH has been implicated in the P450-catalyzed dimerization of ditryptophenalanine and chaetocin.[10,11] This mechanism is supported by the observation that CnsC is inactive toward either N1-methylated indole substrate.

From **14** (or **14**′), the two 3H-indoles in the two halves of the molecule can be subjected to intramolecular nucleophilic attack from the two amine groups and form the pair of aminal linkages. The regioselectivity of the aminal bond formation determines whether the communesin scaffold (as in **4**) or the isocommunesin scaffold (as in **8**) is formed. Our results indicate this regioselectivity is strongly influenced by the C3 substituent of **5**. Energetics of intermediates and products of the reactions starting from **14** or **14**′ were calculated to determine how the additional N-methyl substitution may influence the free energy of aminal C-N bond forming steps (Figure 3). Efforts to locate transition states for these reactions were unsuccessful. We calculated the energies with the N-protonated analogues of these structures to model the protonation states of these species in aqueous solution (Figures S28-S30), but calculations with neutral species yielded

similar results (Figure S27). All intermediates leading to the communesin (**4** and **9**) and isocommunesin scaffolds (**21** and **8**) are thermodynamically accessible at room temperature. The formation of either scaffold is irreversible as scaffold formation is highly exergonic (ca. −25 kcal mol^{-1} for the formation of **4** and **21**, and ca. −28 kcal mol^{-1} for the formation of **8** and **9**). The only energetically uphill step in Figure 3 is the ring opening of pyrroloindole **15** to **16** required for formation of the communesin scaffold. Assuming the activation free energies of each step shown in Figure 3 are proportional to energies for the formation of the intermediates, we would expect that, of the elementary steps shown in Figure 3, the conversion of **15** to **16** would occur most slowly. Thus, formation of the isocommunesin scaffold would be more facile nonenzymatically than formation of the communesin scaffold. Given that the reaction of **5** and **6** with CnsC yields only the communesin product **4**, both the initial oxidative coupling that forms **14** and the subsequent steps involving C-N bond formations and scission must be enzyme-controlled. When the methylated **7** is coupled with **6** to form **14′**, CnsC may no longer be able to exert complete control over the orientation of the intermediate of **14′**, leading to the non-enzymatic formation of **8** and small amount of **9**. Computational analysis also provided explanation for the exclusive formation of the isocommunesin scaffold **11** when **10** was used as a coupling partner to **6**, which is due to a prohibitively high energy intermediate in the reaction path required for the formation of the communesin-like product (Figure S30).

In conclusion, we have identified a P450 monooxygenase CnsC that catalyzes the C3-C3′ coupling of two different indole substrates to introduce vicinal quaternary stereocenters. The enzyme also controls the regioselectivity of the subsequent aminal bond forming steps. The discovery of such a remarkable P450 further illustrates the biocatalytic power of fungal biosynthetic enzymes in the synthesis of complex molecular scaffolds.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b01413.

> Experimental details, spectroscopic and computational data (PDF)

■ AUTHOR INFORMATION

Corresponding Authors
*yitang@ucla.edu
*neilgarg@chem.ucla.edu
*houk@chem.ucla.edu

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) (a) Kim, J.; Movassaghi, M. *Acc. Chem. Res.* **2015**, *48*, 1159. (b) Cherblanc, F. L.; Chapman, K. L.; Brown, R.; Fuchter, M. J. *Nat. Chem. Biol.* **2013**, *9*, 136.

(2) (a) Movassaghi, M.; Schmidt, M. A. *Angew. Chem., Int. Ed.* **2007**, *46*, 3725. (b) May, J. A.; Stoltz, B. M. *Tetrahedron* **2006**, *62*, 5262. (c) Xu, J.-B.; Cheng, K.-J. *Molecules* **2015**, *20*, 6715. (d) Takayama, H.; Matsuda, Y.; Masubuchi, K.; Ishida, A.; Kitajima, M.; Aimi, N. *Tetrahedron* **2004**, *60*, 893. (e) Kitajima, M.; Mori, I.; Arai, K.; Kogure, N.; Takayama, H. *Tetrahedron Lett.* **2006**, *47*, 3199. (f) Büschleb, M.; Dorich, S.; Hanessian, S.; Tao, D.; Schenthal, K. B.; Overman, L. E. *Angew. Chem., Int. Ed.* **2016**, *55*, 2.

(3) (a) Jadulco, R.; Edrada, R. A.; Ebel, R.; Berg, A.; Schaumann, K.; Wray, V.; Steube, K.; Proksch, P. *J. Nat. Prod.* **2004**, *67*, 78. (b) Hayashi, H.; Matsumoto, H.; Akiyama, K. *Biosci., Biotechnol., Biochem.* **2004**, *68*, 753. (c) Andersen, B.; Smedsgaard, J.; Frisvad, J. C. *J. Agric. Food Chem.* **2004**, *52*, 2421. (d) Dalsgaard, P. W.; Blunt, J. W.; Munro, M. H. G.; Frisvad, J. C.; Christophersen, C. *J. Nat. Prod.* **2005**, *68*, 258. (e) Wigley, L. J.; Mantle, P. G.; Perry, D. A. *Phytochemistry* **2006**, *67*, 561. (f) Numata, A.; Takahashi, C.; Ito, Y.; Takada, T.; Kawai, K.; Usami, Y.; Matsumura, E.; Imachi, M.; Ito, T.; Hasegawa, T. *Tetrahedron Lett.* **1993**, *34*, 2355.

(4) (a) Wu, H. X.; Xue, F.; Xiao, X.; Qin, Y. *J. Am. Chem. Soc.* **2010**, *132*, 14052. (b) Yang, J.; Wu, H. X.; Shen, L. Q.; Qin, Y. *J. Am. Chem. Soc.* **2007**, *129*, 13794. (c) Zuo, Z. W.; Xie, W. Q.; Ma, D. W. *J. Am. Chem. Soc.* **2010**, *132*, 13226. (d) Zuo, Z.; Ma, D. *Angew. Chem., Int. Ed.* **2011**, *50*, 12008. (e) Liu, P.; Seo, J. H.; Weinreb, S. M. *Angew. Chem., Int. Ed.* **2010**, *49*, 2000. (f) Seo, J. H.; Liu, P.; Weinreb, S. M. *J. Org. Chem.* **2010**, *75*, 2667. (g) Belmar, J.; Funk, R. L. *J. Am. Chem. Soc.* **2012**, *134*, 16941. (h) Crawley, S. L.; Funk, R. L. *Org. Lett.* **2003**, *5*, 3169. (i) Han, S. J.; Vogt, F.; May, J. A.; Krishnan, S.; Gatti, M.; Virgil, S. C.; Stoltz, B. M. *J. Org. Chem.* **2015**, *80*, 528. (j) Han, S. J.; Vogt, F.; Krishnan, S.; May, J. A.; Gatti, M.; Virgil, S. C.; Stoltz, B. M. *Org. Lett.* **2014**, *16*, 3316. (k) May, J. A.; Zeidan, R. K.; Stoltz, B. M. *Tetrahedron Lett.* **2003**, *44*, 1203. (l) Siengalewicz, P.; Gaich, T.; Mulzer, J. *Angew. Chem., Int. Ed.* **2008**, *47*, 8170.

(5) Lin, H.-C.; Chiou, G.; Chooi, Y.-H.; McMahon, T. C.; Xu, W.; Garg, N. K.; Tang, Y. *Angew. Chem., Int. Ed.* **2015**, *54*, 3004.

(6) Ma, S. M.; Li, J. W. H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X. K.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. *Science* **2009**, *326*, 589.

(7) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. *Nat. Chem. Biol.* **2014**, *10*, 640.

(8) (a) Krishnan, S.; Bagdanoff, J. T.; Ebner, D. C.; Ramtohul, Y. K.; Tambar, U. K.; Stoltz, B. M. *J. Am. Chem. Soc.* **2008**, *130*, 13745. (b) Behenna, D. C.; Krishnan, S.; Stoltz, B. M. *Tetrahedron Lett.* **2011**, *52*, 2152.

(9) Albuquerque, P.; Casadevall, A. *Med. Mycol.* **2012**, *50*, 337.

(10) Saruwatari, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.; Watanabe, K. *ChemBioChem* **2014**, *15*, 656.

(11) (a) Walsh, C. T. *ACS Chem. Biol.* **2014**, *9*, 2718. (b) Gerken, T.; Walsh, C. T. *ChemBioChem* **2013**, *14*, 2256.