VIP Alkaloid Biosynthesis Very Important Paper

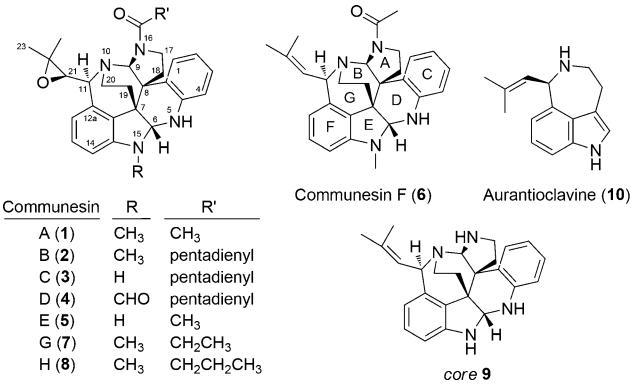
DOI: 10.1002/anie.201411297

Elucidation of the Concise Biosynthetic Pathway of the Communesin Indole Alkaloids**

Hsiao-Ching Lin, Grace Chiou, Yit-Heng Chooi, Travis C. McMahon, Wei Xu, Neil K. Garg, and Yi Tang**

Abstract: *The communesins are a prominent class of indole alkaloids isolated from Penicillium species. Owing to their daunting structural framework and potential as pharmaceuticals, communesins have inspired numerous synthetic studies. However, the genetic and biochemical basis of communesin biosynthesis has remained unexplored. Herein, we report the identification and characterization of the communesin (cns) biosynthetic gene cluster from Penicillium expansum. We confirmed that communesin is biosynthesized by the coupling of tryptamine and aurantioclavine, two building blocks derived from* L*-tryptophan. The postmodification steps were mapped by targeted-gene-deletion experiments and the structural elucidation of intermediates and new analogues. Our studies set the stage for the biochemical characterization of communesin biosynthesis. This knowledge will aid our understanding of how nature generates remarkable structural complexity from simple precursors.*

Many fungal indole alkaloids possess complex structures that are generated in concise biosynthetic pathways.[1] One of the most structurally complex families of indole alkaloids are the communesins (e.g., **1**–**8**, Scheme 1) isolated from marine and terrestrial *Penicillium* species.[2] The common core of the communesins, represented by structure **9**, showcases the intricate molecular architecture of these polycyclic compounds. The natural products contain a total of seven interconnected rings, two aminal linkages, and five (or six) stereocenters. The stereocenters at C7 and C8 are vicinal quaternary centers, which present a considerable synthetic challenge.



Communesin	R	R'
A (**1**)	CH_3	CH_3
B (**2**)	CH_3	pentadienyl
C (**3**)	H	pentadienyl
D (**4**)	CHO	pentadienyl
E (**5**)	H	CH_3
G (**7**)	CH_3	CH_2CH_3
H (**8**)	CH_3	$CH_2CH_2CH_3$

Scheme 1. Structures of the communesins and related compounds.

Because of their remarkable structures, the communesins have inspired numerous synthetic studies,[3] including completed or formal total syntheses by the groups of Qin,[3a,b] Ma,[3c,d] Weinreb,[3e,f] Funk,[3g,h] and Stoltz.[3i–k] In contrast, the biosynthetic pathway of the communesins has remained unexplored since their isolation except for preliminary feeding experiments, which showed that the communesin core is assembled from two indole precursors derived from tryptophan.[2e,4] Stoltz and co-workers proposed several biosynthetic possibilities, all of which involve the coupling of tryptamine and aurantioclavine (**10**) as a key step in the construction of the core structure.[5] To understand how nature generates the structural complexity seen in the communesins, we set out to characterize the responsible alkaloid biosynthetic gene cluster. Herein we report the genetic verification of the *cns* pathway from the known producer *Penicillium expansum* and the isolation of new communesin analogues from blocked mutants.

P. expansum NRRL 976 was confirmed to produce **1** and **2** (Figure 1 B) by comparison of the NMR spectra with reported data (see Tables S3 and S4 and Figures S11–S14 in the Supporting Information).[2b,f,6] Analysis of the sequenced genome by the use of AntiSMASH2[7] revealed that this strain is highly rich in secondary-metabolism gene clusters, including 25 polyketide synthases (PKSs) and 17 nonribosomal peptide synthetases (NRPSs). We reasoned that the *N*-hexadienyl acyl chain in **2** must be derived from a highly reducing PKS, whereas the isoprenoid unit on the indoline must be installed by an enzyme homologous to 4-dimethylallyl tryptophan synthase (DMATS). Furthermore, precursor feeding studies showed that one equivalent of 6-fluorotryptamine can be incorporated into **1** and **2**,[2e,4] thus suggesting

[*] Dr. H.-C. Lin, Dr. W. Xu, Prof. Y. Tang
Department of Chemical and Biomolecular Engineering
Department of Chemistry and Biochemistry
University of California, Los Angeles
Los Angeles, CA 90095 (USA)
E-mail: yitang@ucla.edu
Dr. G. Chiou, Dr. T. C. McMahon, Prof. N. K. Garg
Department of Chemistry and Biochemistry
University of California, Los Angeles
E-mail: neilgarg@chem.ucla.edu
Dr. Y.-H. Chooi
Plant Sciences Division, Research School of Biology
The Australian National University (Australia)

[**] H-C.L. is supported by the National Science Council of Taiwan (102-2917-I-564-008), and G.C. was supported by a NSF Graduate Research Fellowship (DGE-0707424). This research was supported by NIH grants 1DP1GM106413 and 1R56AI101141. We thank Prof. Brian Stoltz for the gift of aurantioclavine. NMR instrumentation is supported by the NSF (CHE-1048804) and the NCRR (S10RR025631).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201411297.

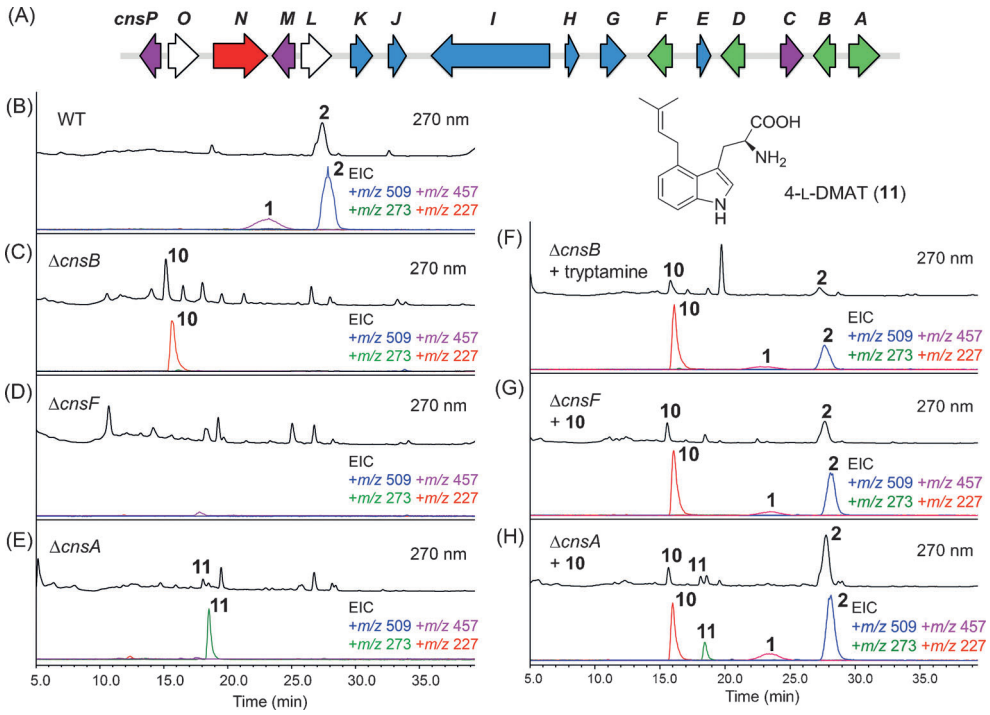


Figure 1. Mapping of the early stages of the communesin biosynthetic pathway. A) The putative *cns* gene cluster from *P. expansum* NRRL 976. Genes are color-coded according to functions of encoded enzymes: precursor synthesis (green), post-core-tailoring modifications (blue), other oxygenase functions (purple), fungal transcription factor (red), and those with unknown/unassigned function (white; see Table S1). B) Confirmation of the production of **1** and **2** from the wild-type (WT) strain. The UV ($\lambda = 270$ nm) trace is shown above, and the selected ion monitoring MS trace is shown below with the masses indicated in different colors. C–E) LC–MS analysis of metabolites extracted from Δ*cnsA*, Δ*cnsB*, and Δ*cnsF*. F–H) LC–MS analysis of metabolites produced from mutant strains shown in (C–E) after chemical complementation with proposed biosynthetic precursors. Tryptamine (F) or **10** (G and H) was supplemented at a concentration of approximately 40 $\mu g\,mL^{-1}$.

the likely presence of a tryptophan decarboxylase (TDC). We therefore searched for a gene cluster that contained a PKS, DMATS, and TDC in close proximity to each other in the *P. expansum* genome.

A candidate gene cluster that encoded these three genes was found in scaffold 19 (Figure 1A; see also Table S2). The putative *cns* gene cluster encodes, besides the PKS (CnsI), DMATS (CnsF), and TDC (CnsB), a methyltransferase (CnsE), an acyltransferase (CnsK), a serine hydrolase (CnsH), and a CoA ligase (CnsG). Furthermore, an assortment of oxidative enzymes are found in the cluster, including a P450 (CnsC), a phytanoyl-CoA dioxygenase (CnsJ), and two non-heme iron-dependent oxygenases (CnsM and CnsP). Interestingly, the gene cluster contains a flavin adenine dinucleotide (FAD)-dependent monooxygenase (CnsA) and a putative catalase (CnsD) that share similarities to chanoclavine-I synthase (EasE) and its catalase partner (EasC), respectively (identity 51 and 59%, respectively).[8] EasC and EasE were reported to be involved in the cyclization of *N*-methyldimethylallyltryptophan (Me-DMAT) to form chanoclavine-I, a key intermediate in the biosynthesis of ergot alkaloids.[9] Hence, CnsA and CnsD may be analogously involved in the synthesis of **10** through the cyclization of 4-DMAT (**11**).

We first inactivated *cns* genes that may be involved in precursor synthesis, such as *cnsA*, *cnsB* (TDC), and *cnsF* (DMATS), through double-crossover recombination by using the hygromycin-resistance gene *hyg* as a marker (see Figures S1–S8). In each knockout strain, the biosynthesis of both **1** and **2** was abolished, thus confirming the essential roles of these enzymes and the *cns* cluster in communesin biosynthesis (Figure 1C–E). The Δ*cnsB* mutant strain showed the accumulation of a new compound with $\lambda_{max} = 286$ nm and *m/z* 227 $[M+H]^+$, which was characterized as aurantioclavine (**10**) by comparison with an authentic sample (see Table S7 and Figures S19 and S20). When the Δ*cnsB* strain was complemented with tryptamine, the biosynthesis of **1** and **2** was restored (Figure 1F), thus confirming the role of CnsB in the generation of tryptamine.

In contrast, deletion of *cnsF* did not lead to the appearance of compounds that may be involved in the communesin pathway (Figure 1D). We hypothesized that this mutant cannot produce 4-DMAT (**11**) and subsequently **10**. The absence of tryptamine in the extract may be due to fungal metabolic consumption. Chemical complementation of Δ*cnsF* with **10** restored the levels of **1** and **2** (Figure 1G), thus confirming that **10** is indeed a true intermediate en route to communesins, as proposed by Stoltz and co-workers.[3k] To assay the substrate specificity of CnsF towards either L-Trp or tryptamine, the recombinant protein was expressed and purified from *Escherichia coli* BL21 (DE3) (see Figure S9). CnsF displayed much higher turnover in converting L-Trp entirely into **11**. In contrast, only partial conversion of tryptamine into 4-dimethylallyltryptamine was observed (see Figure S10). This result is consistent with the assignment of CnsF as 4-DMATS on the basis of its sequence similarity to FgaPT2.[10]

The Δ*cnsA* mutant strain accumulated a compound with $\lambda_{max} = 273$ nm and *m/z* 273 $[M+H]^+$, which was verified to be **11** (see Table S8 and Figure S21). Complementation of this strain with **10** restored the production of **1** and **2** (Figure 1H). Hence, CnsA was assigned as the aurantioclavine synthase that catalyzes the conversion of **11** into **10**. This conversion probably starts with the flavin-dependent oxidation of the benzylic carbon atom in **11**, followed by intramolecular attack by the α-NH_2 group. The subsequent decarboxylation then leads to the formation of **10** (Scheme 2). This conversion may be aided similarly by the catalase CnsD, as is the case during

chanoclavine biosynthesis.[8b] Taken together, our knockout and chemical-complementation experiments revealed the divergent synthesis of two different indole-containing building blocks, tryptamine and **10**, from L-Trp.

We then turned to the downstream enzymes that may modify the core **9** through methylation, acylation, and epoxidation. The N15-methylation and N16-acylation reactions take place on the secondary amines on opposite poles of the core and may be important in stabilizing the otherwise labile aminal linkages. Deletion of *cnsE*, which encodes the sole methyltransferase in the cluster, led to the disappearance of **1** and **2**, and the emergence of two compounds **3** (λ_{max} = 268 nm, *m/z* 495 $[M+H]^+$) and **5** (λ_{max} = 269 nm, *m/z* 443 $[M+H]^+$), which are the desmethyl versions of **2** and **1**, respectively (Figure 2A). The structure of **3** was confirmed by NMR spectroscopy to be communesin C (see Table S5 and Figures S15 and S16).

The co-isolation of **1** and **2** suggests that N16-acylation by acetyl-CoA and the hexadienyl unit should be in competition. Acylation to introduce the hexadienyl chain in **2** is reminiscent of the acylation of fumagillol by a pentaene polyketide in the biosynthesis of fumagillin.[11] In the *cns* pathway, *cnsI* encodes a highly reducing PKS that has ketoreductase and dehydratase domains, but no enoylreductase domain. As expected, knockout of *cnsI* led to the exclusive formation of **1** (Figure 2B). The gene *cnsK* encodes an enzyme belonging to the transferase family (PFAM02458), and may play a role in transferring the polyketide chain to N16 of the core.

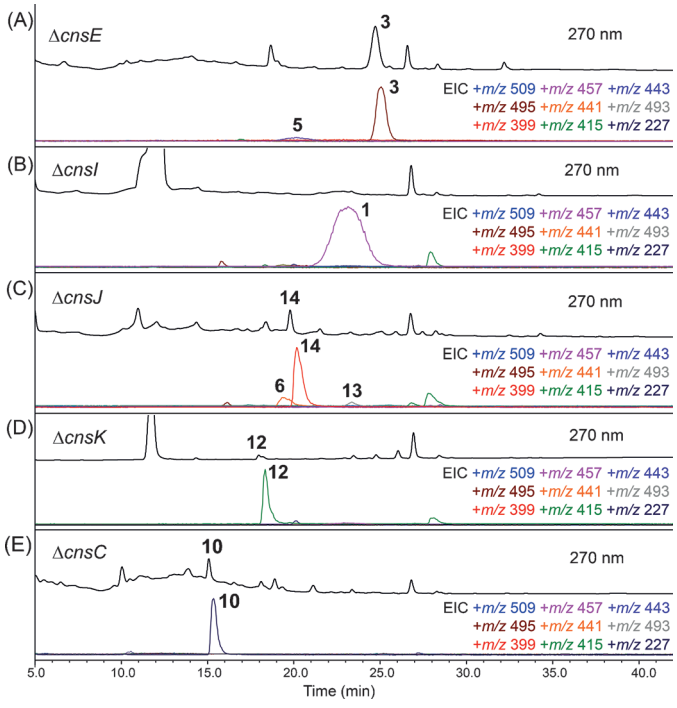
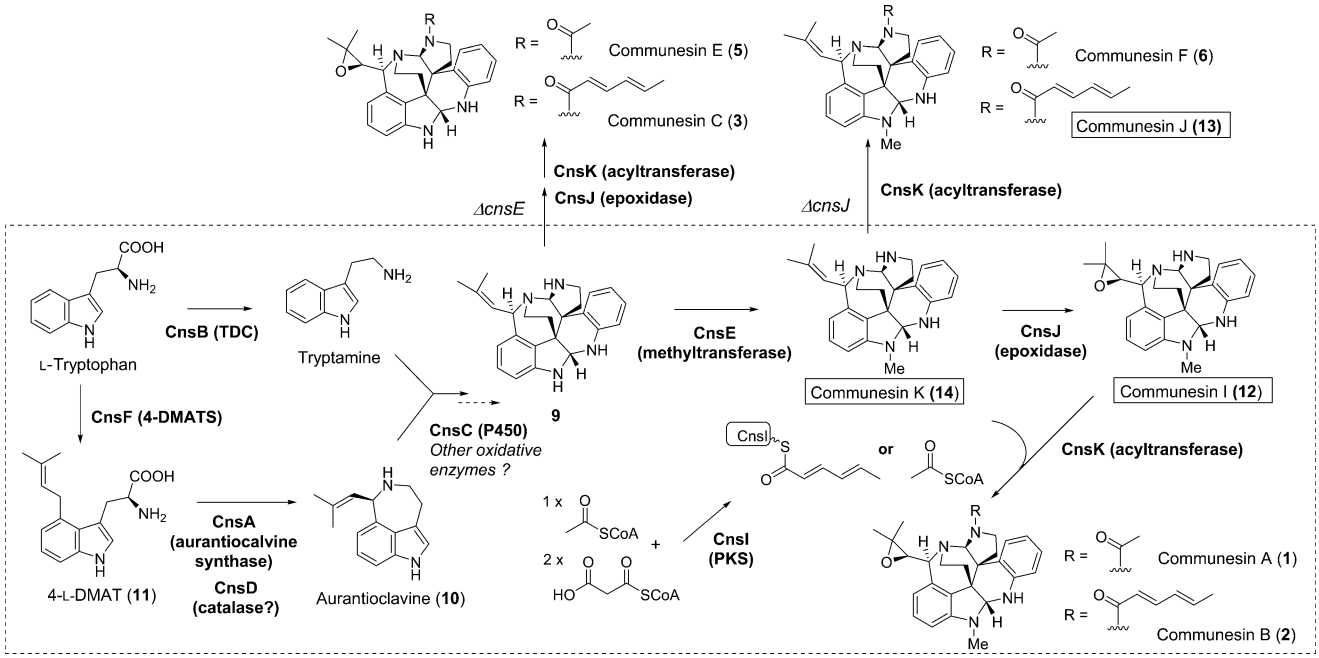


Figure 2. Mapping of the late stages of the communesin biosynthetic pathway. Confirmation of the roles of enzymes encoded by A) *cnsE* (methyltransferase), B) *cnsI* (PKS), C) *cnsJ* (epoxidase), and D) *cnsK* (acyltransferase). E) The role of the enzyme encoded by *cnsC* (P450) was also investigated. In each case, the selected ion monitoring MS trace with the masses indicated in different colors is shown below the UV trace (λ = 270 nm).

Deletion of *cnsK* abolished the production of both **1** and **2**. Instead, a single product **12** (λ_{max} = 267 nm, *m/z* 415 $[M+H]^+$) was produced (Figure 2D). The new compound was structurally characterized as the non-acylated version of **1** and **2** (see Table S9 and Figures S22 and S23). Production of **12** (communesin I) confirmed the role of CnsK as the N16 acyltransferase. It is possible that the hexadienyl product is first hydrolytically removed from the PKS by the serine hydrolase CnsH, converted into hexadienyl-CoA by the CoA ligase CnsG, and then transferred to **12** by CnsK. Surprisingly, the simultaneous loss of **1** in this mutant strain suggests that CnsK may also be a promiscuous acyltransferase that can tolerate a range of acyl groups, including acetyl-, propionyl-, and butyryl-CoA, which lead to **1**, **7**, and **8**, respectively. Exploration of different acyl groups could lead to the production of additional non-natural communesin analogues.

The gene *cnsJ* encodes a putative oxidative enzyme that displays 40% sequence similarity to FtmF found in the fumitremorgin pathway.[12] CnsJ may therefore install the C21–C22 epoxide in the prenylated portion of **1** and **2**. Indeed, deletion of *cnsJ* led to the production of **6** (*m/z* 441 $[M+H]^+$), **13** (*m/z* 493 $[M+H]^+$), and **14** (*m/z* 399 $[M+H]^+$) (Figure 2C), all of which were structurally characterized (see Tables S6, S10, and S11 and Figures S24–S32). The predominant product **14** (communesin K) is a new compound that is not acylated at N16 and non-oxidized at C21–C22. The two minor products **6** and **13** are communesin F and a new analogue (communesin J), respectively (Scheme 2). The accumulation of non-acylated **14** as the major product suggests that the downstream acyl-transfer reaction catalyzed by CnsK may be sluggish in the absence of the C21–C22 epoxide and most likely takes place after the actions of CnsJ (Scheme 2). Communesin K (**14**) is the simplest characterized communesin that contains the heptacyclic core. We also constructed a Δ*cnsJ*/Δ*cnsE* double mutant in an attempt to isolate **9**. Although selective ion monitoring showed the presence of a new compound with UV and mass spectra consistent with those of **9**, its isolation was unsuccessful owing to rapid degradation of the molecule. Although we could not pinpoint the timing of the methylation at N15 with the knockout studies, the structure of **14** coupled with its significantly improved stability as compared to that of **9** suggests that the methylation step takes place immediately after the formation of **9** to yield **14**.

Oxidative coupling of **10** and tryptamine was proposed to be a possible route to the formation of the challenging vicinal quaternary centers of the communesins.[5] We therefore investigated the role of CnsC, a cytochrome P450 encoded in the cluster. Protein-sequence analysis suggests CnsC may be an atypical P450 with low sequence homology (< 25%) to those in the P450 database.[13] Genetic inactivation of *cnsC* led to the abolishment of **1** and **2**, and the accumulation of **10** (Figure 2E). Since tryptamine was not detectable in the culture, we supplied tryptamine to *ΔcnsC* to ensure that the block mutant was not defective in tryptamine biosynthesis. Supplementation of both tryptamine and **10** did not restore communesin biosynthesis. These results therefore hint that CnsC acts



Scheme 2. Biosynthetic pathway of the communesins as mapped on the basis of the knockout studies presented herein. The pathway in the dashed box is the main pathway leading to **1** and **2** in *P. expansum*. The compounds outside of the box were isolated as shunt products by the use of blocked mutants. Three new communesins, **12**, **13**, and **14**, were produced.

downstream of both **10** and tryptamine, and probably catalyzes a key step in the coupling of these two indole derivatives to form **9**. The exact roles of this enzyme, along with other oxidative enzymes, such as CnsP, are currently under investigation by the use of biochemical means.

In conclusion, the communesin biosynthetic pathway was mapped by genetic-inactivation studies, which led to the isolation of three new communesin analogues. The involvement of aurantioclavine as a building block was confirmed. Our studies provide a solution to the biosynthetic mystery surrounding this famous class of indole alkaloids and set the stage for detailed biochemical characterization to understand how nature generates exceptional structural complexity from simple building blocks.

Received: November 24, 2014
Published online: January 8, 2015

Keywords: biosynthesis · communesins · gene deletion · indole alkaloids · structural elucidation

[1] W. Xu, D. J. Gavia, Y. Tang, *Nat. Prod. Rep.* **2014**, *31*, 1474 – 1487.
[2] a) R. Jadulco, R. A. Edrada, R. Ebel, A. Berg, K. Schaumann, V. Wray, K. Steube, P. Proksch, *J. Nat. Prod.* **2004**, *67*, 78 – 81; b) H. Hayashi, H. Matsumoto, K. Akiyama, *Biosci. Biotechnol. Biochem.* **2004**, *68*, 753 – 756; c) B. Andersen, J. Smedsgaard, J. C. Frisvad, *J. Agric. Food Chem.* **2004**, *52*, 2421 – 2428; d) P. W. Dalsgaard, J. W. Blunt, M. H. G. Munro, J. C. Frisvad, C. Christophersen, *J. Nat. Prod.* **2005**, *68*, 258 – 261; e) L. J. Wigley, P. G. Mantle, D. A. Perry, *Phytochemistry* **2006**, *67*, 561 – 569; f) A. Numata, C. Takahashi, Y. Ito, T. Takada, K. Kawai, Y. Usami, E. Matsumura, M. Imachi, T. Ito, T. Hasegawa, *Tetrahedron Lett.* **1993**, *34*, 2355 – 2358.
[3] a) H. Wu, F. Xue, X. Xiao, Y. Qin, *J. Am. Chem. Soc.* **2010**, *132*, 14052 – 14054; b) J. Yang, H. Wu, L. Shen, Y. Qin, *J. Am. Chem. Soc.* **2007**, *129*, 13794 – 13795; c) Z. Zuo, W. Xie, D. Ma, *J. Am. Chem. Soc.* **2010**, *132*, 13226 – 13228; d) Z. Zuo, D. Ma, *Angew. Chem. Int. Ed.* **2011**, *50*, 12008 – 12011; *Angew. Chem.* **2011**, *123*, 12214 – 12217; e) P. Liu, J. H. Seo, S. M. Weinreb, *Angew. Chem. Int. Ed.* **2010**, *49*, 2000 – 2003; *Angew. Chem.* **2010**, *122*, 2044 – 2047; f) J. H. Seo, P. Liu, S. M. Weinreb, *J. Org. Chem.* **2010**, *75*, 2667 – 2680; g) J. Belmar, R. L. Funk, *J. Am. Chem. Soc.* **2012**, *134*, 16941 – 16943; h) S. L. Crawley, R. L. Funk, *Org. Lett.* **2003**, *5*, 3169 – 3171; i) S.-J. Han, F. Vogt, J. A. May, S. Krishnan, M. Gatti, S. C. Virgil, B. M. Stoltz, *J. Org. Chem.* **2014**, *80*, 528 – 547; j) S.-J. Han, F. Vogt, S. Krishnan, J. A. May, M. Gatti, S. C. Virgil, B. M. Stoltz, *Org. Lett.* **2014**, *16*, 3316 – 3319; k) J. A. May, R. K. Zeidan, B. M. Stoltz, *Tetrahedron Lett.* **2003**, *44*, 1203 – 1205; l) P. Siengalewicz, T. Gaich, J. Mulzer, *Angew. Chem. Int. Ed.* **2008**, *47*, 8170 – 8176; *Angew. Chem.* **2008**, *120*, 8290 – 8296.
[4] L. J. Wigley, D. A. Perry, P. G. Mantle, *Mycol. Res.* **2008**, *112*, 131 – 137.
[5] J. A. May, B. Stoltz, *Tetrahedron* **2006**, *62*, 5262 – 5271.
[6] I. Kerzaon, Y. F. Pouchus, F. Monteau, B. Le Bizec, M.-R. Nourrisson, J.-F. Biard, O. Grovel, *Rapid Commun. Mass Spectrom.* **2009**, *23*, 3928 – 3938.
[7] K. Blin, M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano, T. Weber, *Nucleic Acids Res.* **2013**, *41*, W204 – W212.
[8] a) N. Lorenz, J. Olšovská, M. Šulc, P. Tudzynski, *Appl. Environ. Microbiol.* **2010**, *76*, 1822 – 1830; b) K. E. Goetz, C. M. Coyle, J. Z. Cheng, S. E. O'Connor, D. G. Panaccione, *Curr. Genet.* **2011**, *57*, 201 – 211.
[9] C. A. F. Nielsen, C. Folly, A. Hatsch, A. Molt, H. Schröder, S. E. O'Connor, M. Naesby, *Microb. Cell Fact.* **2014**, *13*, 95.
[10] I. A. Unsöld, S.-M. Li, *Microbiology* **2005**, *151*, 1499 – 1505.
[11] H.-C. Lin, Y.-H. Chooi, S. Dhingra, W. Xu, A. M. Calvo, Y. Tang, *J. Am. Chem. Soc.* **2013**, *135*, 4616 – 4619.
[12] N. Steffan, A. Grundmann, S. Afiyatullov, H. Ruan, S.-M. Li, *Org. Biomol. Chem.* **2009**, *7*, 4082 – 4087.
[13] D. R. Nelson, *Hum. Genomics* **2009**, *4*, 59 – 65.

