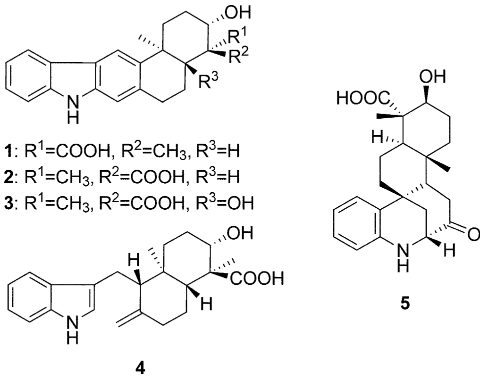
DOI: 10.1002/anie.201204087

Bacterial Synthesis of Diverse Indole Terpene Alkaloids by an Unparalleled Cyclization Sequence**

*Zhongli Xu, Martin Baunach, Ling Ding, and Christian Hertweck**

Indole terpenoids encompass a highly diverse group of natural products, including infamous psychotropic agents such as lysergic acid derivatives, the aphrodisiac yohimbine, and the potassium channel blockers paxilline and lolitrem.[1] What is remarkable about this multifarious class is that practically all indole terpene alkaloids have been isolated from plants and fungi.[2] Stimulated by the importance of these synthetically challenging compounds, groundbreaking studies have been undertaken in recent years to understand and engineer terpene alkaloid pathways in fungi and plants.[3] In light of the impressive number of known eukaryotic indole terpene metabolites, it is peculiar that only recently the first bacterial representatives of this group were discovered. We and others independently reported the structures of pentacyclic indolocarbazoles from *Streptomyces* spp., namely the diastereomers oridamycin (**1**)[4] and xiamycin A (**2**; Scheme 1).[5] Considering that these indolosesquiterpenes (IST) are reminiscent of plant metabolites, it is astounding that two xiamycin-producing strains are endophytes of widespread mangrove trees, *Bruguiera gymnorrhiza*[5] and *Kandelia candel*.[6] A more detailed metabolic investigation of the *K. candel* endophyte revealed three congeners of **2**, xiamycin B (**3**), the seco-derivative indosespene (**4**), and the novel bridged spiro compound sespenine (**5**).[6] These rare endophyte metabolites likely play an ecological role in their habitats because their diverse antiviral, antibacterial, and antifungal activities may contribute to the antibiotic reservoir of the mangrove plants.[7] From a chemical point of view, the co-occurrence of these structurally novel alkaloids is intriguing because it suggests a common biogenetic origin. However, to date nothing is known about the biosynthesis of indole terpenes in bacteria. Herein we unveil the molecular basis for unprecedented bacterial indolosesquiterpene biosynthesis in a mangrove endophyte and show by heterologous gene expression and mutational analysis that the unusual pentacyclic ring systems of xiamycin and sespenine are formed by a novel cyclization sequence. Furthermore, we report the discovery of three new xiamycin dimers from a heterologously reconstituted IST pathway.



Scheme 1. Structures of bacterial indole sesquiterpenes oridamycin (**1**), xiamycin A (**2**) and B (**3**), indosespene (**4**), and sespenine (**5**).

To elucidate the genes required for IST biosynthesis in the mangrove endophyte, we subjected whole genomic DNA to shotgun sequencing. Bioinformatic mining (basic local alignment search tool (BLAST) analysis) of the genomic draft sequence of *Streptomyces* sp. HKI0576 for terpenoid biosynthesis genes revealed a gene cluster (GenBank accession No. HE815466) with a three-gene cassette (*xiaABC*) coding for canonical enzymes involved in the non-mevalonate (or deoxyxylulose, DOX) pathway:[8] DXS (1-deoxy-D-xylulose 5-phosphate synthase), HDS (4-hydroxy-3-methylbut-2-enyl diphosphate synthase) and HDR (4-hydroxy-3-methylbut-2-enyl diphosphate reductase); remaining DOX pathway genes were identified elsewhere in the genome. The isoprene unit biosynthesis genes are flanked by genes coding for regulatory components, two putative polyprenyl synthetases, and various oxidoreductases. An important clue in the identification of the biosynthetic gene cluster was the finding of a putative indole oxygenase gene, which suggested that this gene cluster could play a role in indole terpenoid biosynthesis. We could verify this assumption through various lines of evidence, specifically by targeted gene deletions and heterologous expression of the entire gene locus. First, we selected a cosmid (04B02) harboring the predicted *xia* biosynthesis gene cluster (Figure 1) and subcloned the insert (ca. 38 kb) into a *Streptomyces–E. coli* shuttle vector (pKJ55).[9] The resulting construct, pXU472, was introduced into the heterologous host *Streptomyces albus*. By HPLC-HRMS (Exactive) monitoring we could detect **2** ($C_{23}H_{25}NO_3$, *m*/*z* 362.176 $[M-H]^-$), **3**

---

[*] Dr. Z. Xu, M. Baunach, Dr. L. Ding, Prof. Dr. C. Hertweck
Dept. of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology, HKI
Beutenbergstrasse 11a, 07745 Jena (Germany)
E-mail: christian.hertweck@hki-jena.de
Homepage: http://www.hki-jena.de
Prof. Dr. C. Hertweck
Chair of Natural Product Chemistry, Friedrich Schiller University, Jena (Germany)

[**] Z.X. and M.B. contributed equally to this work. Financial support by the BMBF (GenBioCom) is gratefully acknowledged. We thank W. Karguth for assistance in molecular studies, A. Perner and H. Heineke for MS and NMR measurements, and M. Roth and M. Steinacker for support with fermentations.

Supporting information for this article, including experimental details and a table of deduced *xia* gene functions, is available on the WWW under http://dx.doi.org/10.1002/anie.201204087.

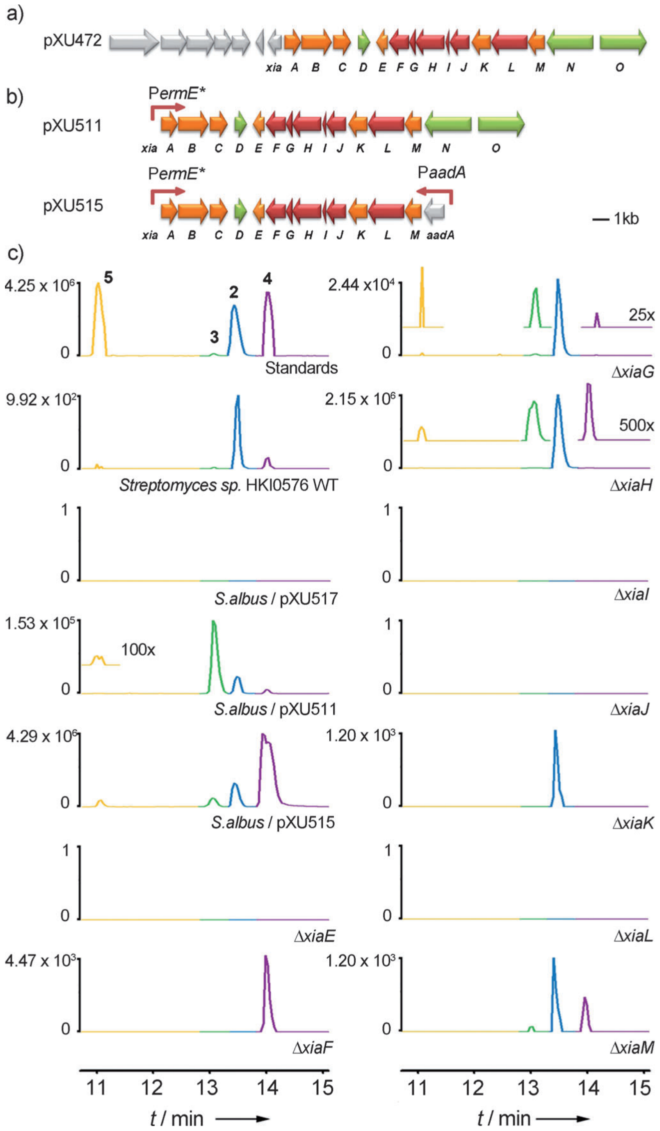
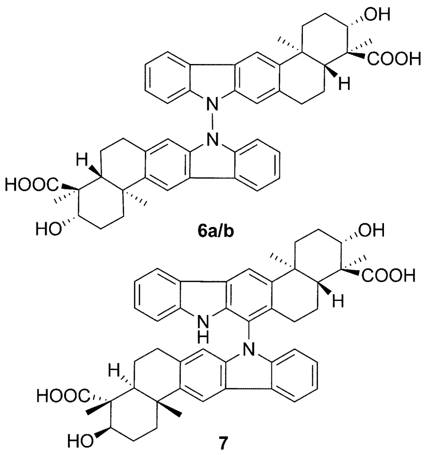


Figure 1. a) Organization of the xiamycin biosynthesis gene cluster in *Streptomyces* sp. HKI0576, b) architecture of expression constructs, and c) results from mutational analyses. Condensed results of HPLC-HRMS monitoring (Exactive) of targeted mutants using single-ion monitoring (magnification offset). Deletions generated on the chromosome of IST wild-type producer *S.* sp. HKI0576; pXU517, blank vector as negative control. The *y* axis represents the relative intensity of the ion current.

($C_{23}H_{25}NO_4$, *m/z* 378.171 $[M-H]^-$), **4** ($C_{23}H_{29}NO_3$, *m/z* 366.207 $[M-H]^-$), and **5** ($C_{23}H_{29}NO_4$, *m/z* 384.217 $[M+H]^+$), showing that genes located on the cosmid insert confer the ability to produce xiamycin and all congeners. Next, we determined the exact boundaries of the gene cluster through λ-Red-mediated[10] truncation of the cosmid insert. In this way, a 21.3 kb construct was generated (pXU511) that comprises solely the *xia* gene locus (*xiaA*-*xiaO*). Again, the expression plasmid was introduced into *S. albus*, and metabolite analysis revealed that all four ISTs are produced at the wild-type level. To increase the production and to facilitate the detection of IST, transcriptional regulator genes (*xiaN*, *xiaO*) upstream of the *xiaM-E* operon were replaced by a resistance cassette with constitutive promoter. We noted that the promoter of *aadA* (the spectinomycin resistance gene) is able to read through the whole *xiaM-E* operon in plasmid pXU515, whereas *xiaA-D* is under control of the promoter PermE*. This set-up resulted in the constitutive expression of all *xia* genes and provided substantially increased titers of **2**–**5**, albeit at a different ratio compared to the wild type. In summary, these experiments not only demonstrated that **2**–**5** indeed share a common biosynthetic origin, but also unequivocally showed that the cloned *xia* genes are essential and sufficient for their biosynthesis.

The heterologous expression of the *xia* genes allowed for the detection and isolation of previously unknown congeners that are somewhat reminiscent of dimers of the pentacyclic polyphenol benastatin.[11] We succeeded in isolating small amounts of **6a**, **6b**, and **7** from polar fractions (Scheme 2).

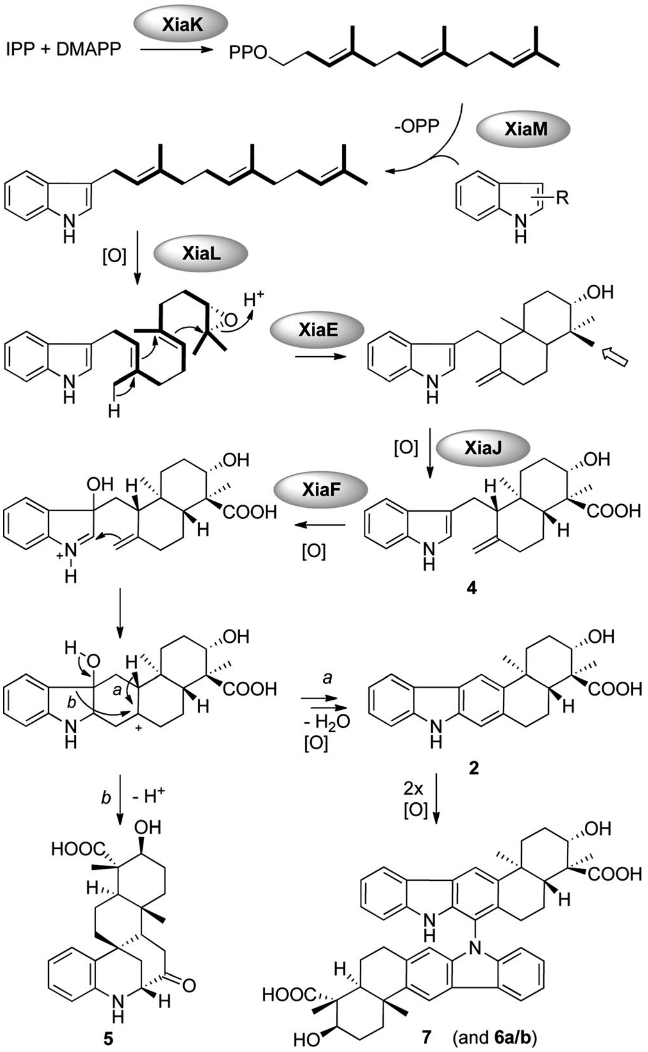


Scheme 2. Structures of C–N- and N–N-linked xiamycin dimers.

The NMR spectra of **6a** and **6b** indicated that they have the xiamycin A backbone, yet according to HRESIMS data their molecular formula is $C_{46}H_{48}N_2O_6$. Consequently, **6a** and **6b** are dimers of xiamycin. Although there are certain possibilities of dimerization, all possibilities except N–N bond formation would give rise to dramatic changes in the chemical shifts in the ^{13}CNMR spectra. Considering the fact that all ^{13}CNMR data were conserved, only the possibility of a N–N dimer was plausible. Owing to the lack of rotational freedom of the two carbazole moieties, **6a** and **6b** were proposed to be a pair of atropoisomers. Compound **7** had the molecular composition of $C_{46}H_{48}N_2O_6$ according to HRESIMS, but in contrast to **6a** and **6b** the ^{1}HNMR spectrum of **7** showed two sets of signals, indicating that the two halves of the molecule were not identical. As only a signal for 21-H ($\delta = 6.53$) was visible, whereas the counterpart of 21′-H was missing in the ^{1}HNMR spectrum, C–N bond formation between N1 and C21′ was proposed.

With the *xia* genes at hand, we next sought to dissect the pathway and functionally characterize the *xia* genes. For this purpose we performed bioinformatic analyses and individually knocked out all crucial pathway genes. BLAST searches revealed that two polyprenyl diphosphate synthetases, XiaK and XiaM, are encoded in the *xia* gene cluster. Both deduced gene products exhibit the universally conserved domain, a highly acidic DDXXD motif, for binding the essential divalent cation during the elongation and/or cyclization of the isoprenoid precursor. However, XiaM is evolutionarily more distant from canonical polyprenyl synthetases, suggesting that it could mediate prenylation of an indole nucleophile in lieu of a terpene building block. We tested this assumption using directed gene knock-outs and, in fact, the main metabolite, **2**, can be detected in both *ΔxiaK* and *ΔxiaM* mutants, albeit only in low amounts. It appears that both enzymes can cross-complement each other to some degree, and bacteria may complement a lacking FPP synthase by recruiting the enzyme or substrate from primary metabolism. However, XiaM lacks the IPP binding site and thus appears to be the first specific enzyme in the pathway. The different metabolic profiles of the *ΔxiaK* and *ΔxiaM* mutants most likely results from the introduction of a strong promoter by the λ RED-mediated gene replacements, which leads to an increased expression of downstream genes.

Another interesting finding was that *xiaL* codes for an unprecedented hybrid protein. Homology searches revealed that the C-terminus (400 aa) is closely related to various FAD-dependent epoxidases involved in the biosynthesis of polyethers and long-chain polyenes, such as monensin,[12] nigericin,[13] and ambruticin,[14] and other enzymes belonging to the superfamily of 2-polyprenyl-6-methoxyphenol hydroxylases (UbiH). However, XiaL is unique, as the second domain (180 aa) is homologous to the N-terminus of tryptophan 2,3-dioxygenases. It was shown that the N-terminus is crucial for indole binding, and an aa alignment showed that all residues required for indole binding (Phe 51, Tyr 24, Tyr 27, Leu 28, His 55)[15] are conserved in XiaL. Thus, we propose that XiaL represents a monooxygenase for long-chain polyenes/terpenes that is fine-tuned for an indole derivative. To test the role of XiaL, we performed directed homologous recombination and obtained a *ΔxiaL* mutant. As expected, none of the known ISTs could be detected in the culture broth of the *ΔxiaL* mutant. The tentative indolylfarnesyl epoxide intermediate (Scheme 3) would be the predicted substrate of a novel type of terpene cyclase, XiaE. When deleting *xiaE* from the gene cluster, we found that the *xia* pathway is fully abrogated in the mutant and concluded that XiaE is indispensable for IST biosynthesis. Initially, a BLAST analysis annotated the gene product of *xiaE* as an unknown membrane protein. Interestingly, *xiaE* homologues are found in various fungal and bacterial terpenoid biosynthesis gene clusters,[16] for example for paspaline biosynthesis;[17] however, their true function has been overlooked or misinterpreted. A more detailed domain analysis indicated that XiaE is functionally related to a recently described homologue that plays a key role in the biosynthesis of the fungal meroterpenoid pyripyropene A. Through stepwise reconstruction of the enzymes involved and upon the feeding of precursors, this enzyme was unequivocally confirmed to be a novel membrane-bound terpene cyclase.[18] The cyclization is initiated by the protonation of the terminal epoxide ring and followed by electrophilic cascade reactions. A phylogenetic analysis revealed that XiaE is uniquely positioned between clades of enzymes encoded in fungal indole terpenoid and meroterpenoid biosynthesis gene clusters (Figure 2).



Scheme 3. Model of indolosesquiterpene biosynthesis in the mangrove endophyte *Streptomyces* sp. HKI0576.

Notably, except for the cyclase from the pyripyropene cyclase, none of these deduced gene products have been characterized. Furthermore, we found a single bacterial relative that is encoded in the biosynthesis gene cluster for the bacterial meroterpenoid BE-40644,[19] but its function has not been elucidated. Thus, XiaE is the first bacterial representative of this novel type of terpene cyclases and the first functionally characterized member of this family that plays a key role in indole terpenoid cyclization.

In this context, it should also be considered that en route to **4**, one isoprenoid methyl group undergoes a six-electron oxidation to yield a carboxy group. We reasoned that XiaJ could catalyze this reaction because it has similarity to terpene-modifying oxygenases, such as linalool 8-monooxygenases. XiaI might represent the corresponding ferredoxin, and indeed both are crucial for indosespene formation, as its formation was completely abolished in both the *ΔxiaJ* and *ΔxiaI* mutants (Figure 1). Another role for XiaJ could be the

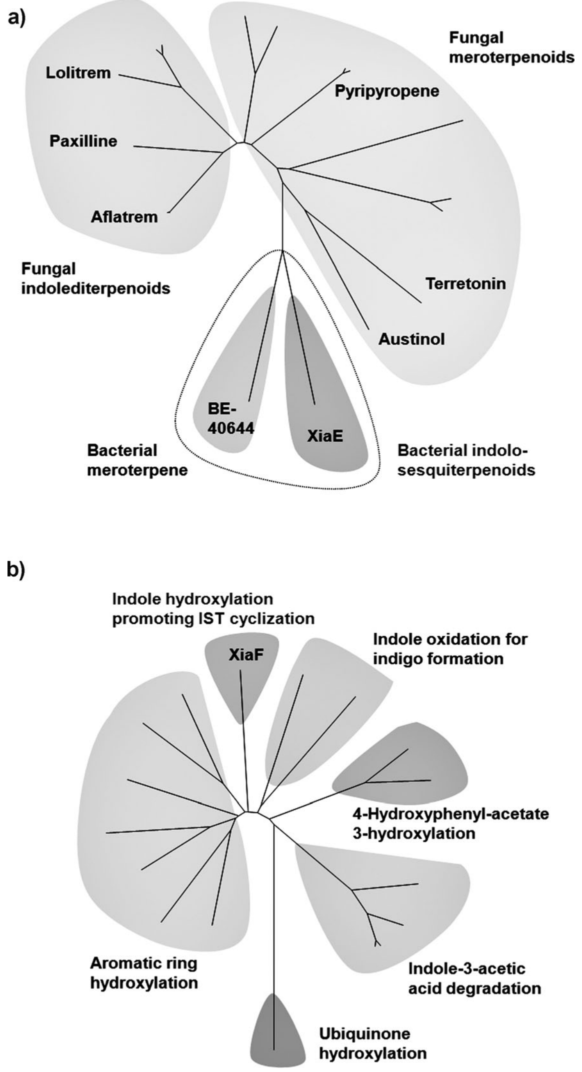


Figure 2. Phylogenetic analyses of a) XiaE and b) XiaF, which are key enzymes for indolosesquiterpene cyclization (for detailed cladograms, see the Supporting Information).

hydroxylation of xiamycin A to yield xiamycin B. Highly sensitive HPLC-HRMS analyses of the mutants indicated the presence of various potential precursors, including the hydrolysis product of the tentative indolylfarnesyl epoxide. However, their scarcity and amphiphilic nature has hampered their isolation and full characterization.

To solve the riddle of how **4** is transformed into the pentacyclic downstream products, we analyzed the remaining candidate genes. On the basis of mechanistic considerations and synthetic model reactions,[20] we have previously proposed that **2** and **5** would emerge from **4** through indole oxygenation with subsequent cyclization/rearrangement.[6] Indeed we noticed a gene (*xiaF*) coding for a putative indole oxygenase that belongs to the superfamily of acyl-CoA dehydrogenases (ACAD). To establish the role of XiaF for xiamycin and sespenine biosynthesis, we generated a targeted gene knockout. Through metabolic profiling we found that the mutant completely lost the ability to produce **2**, **3** and **5**, but accumulated **4** instead (Figure 1). This clear result, together with the annotation, indicates that XiaF mediates a key cyclization step in the *xia* pathway. What is remarkable about this finding is that enzymes related to XiaF have been characterized as indole oxygenases or "pigment producing hydroxylases" because of their ability of generating indoxyl units, which yield indigo and related condensation products.[21] Whereas most members of this group have been shown to promote indigo biosynthesis when heterologously produced in *E. coli*, this is often not their natural native function. Their true substrates are alternative carbon sources, such as aromatics (naphthalene, biphenol) or terpenes (limonene).[22] A phylogenetic analysis of XiaF and its homologues (Figure 2) revealed that XiaF has a particular position between clades of oxygenases involved in indigo formation and aromatics-decomposing relatives. Furthermore, as the three-gene cassette *xiaF-H* shows a marked similarity to clustered orthologues for xenobiotic degradation[23] in various bacteria, it can be speculated that *xiaF-H* was acquired by horizontal gene transfer and has since undergone a change of function. Notably, whereas the loss of XiaG and XiaH has no significant influence on IST formation, XiaF is totally essential for the formation of the pentacyclic metabolites.

Interestingly, this is the first case where an indole oxygenase initiates the cyclization of an indole terpenoid, and this mechanism is fully unprecedented for indole terpenoid pathways. A plausible explanation for the diverging route to the distinctive skeletons of **2** and **5** is the degree of freedom in the rearrangement of the carbenium intermediate. Whereas xiamycin results from cyclization and oxidative aromatization, sespenine formation involves phenyl migration to the angular carbenium ion. The latter route entangles the disruption of the indole ring with concomitant formation of the bridged spiro skeleton of **5** (Scheme 3). Future biochemical and structural studies will shed light on the detailed mechanism of the cyclization processes within the enzyme pocket.

During the past years, genome analyses have propelled the discovery of bacterial terpene metabolites and the corresponding genes.[24] In this study, we have unveiled the molecular basis for the first indole terpenoid pathway in a bacterium. Through heterologous reconstitution of the xiamycin pathway we elucidated the set of *xia* genes that are sufficient for IST biosynthesis and demonstrated the biogenetic relationship of the diverse alkaloids. Furthermore, to our knowledge, this is the first heterologous expression of an intact pathway of an endophyte. As genetic engineering of the *xia* gene cluster resulted in increased metabolite production, we succeeded in the characterization of yet overlooked, structurally unique N−N- and C−N-coupled xiamycin dimers. Through computational studies, mutational analysis of all pathway genes, and metabolite analyses, we succeeded in dissecting an unparalleled bacterial alkaloid pathway. Our results reveal the involvement of two polyprenyl synthetases, XiaK and XiaM, for FPP formation and indole prenylation, and suggest concomitant epoxidation by a novel hybrid enzyme, XiaL. On biosynthetic grounds the most important discoveries were two novel cyclases, which play key roles in indole terpenoid cyclization: XiaE, which promotes the cyclization of the epoxidized farnesyl side chain, and the indole oxygenase XiaF, which paves the way for diverging cyclization. Because of the unusual IST architectures and the novel mechanisms involved, the *xia* biosynthesis enzymes are an important addition to the emerging family of bacterial terpene synthases and modifying enzymes. Thus, this study not only sets the stage for studying mechanistically intriguing reactions, but also is a prerequisite for engineering indoloterpenoids in bacteria.

Received: May 25, 2012
Revised: July 25, 2012

Keywords: alkaloids · indole oxygenase · indolosesquiterpenes · terpene cyclase · xiamycin

---

[1] W. Steglich, B. Fugmann, S. Lang-Fugmann, *Römpp Encyclopedia of Natural Products*, Georg Thieme, Stuttgart, **2002**.

[2] a) S. E. O'Connor, E. McCoy in *Recent Advances in Phytochemistry*, *Vol. 40* (Ed.: J. T. Romeo), Elsevier, **2006**; b) S. Saikia, M. J. Nicholson, C. Young, E. J. Parker, B. Scott, *Mycol. Res.* **2008**, *112*, 184–199.

[3] a) A. R. Usera, S. E. O'Connor, *Curr. Opin. Chem. Biol.* **2009**, *13*, 492–498; b) W. Runguphan, S. E. O'Connor, *Nat. Chem. Biol.* **2009**, *5*, 151–153; c) W. Runguphan, X. Qu, S. E. O'Connor, *Nature* **2010**, *468*, 461–464; d) S. M. Li, *Nat. Prod. Rep.* **2010**, *27*, 57–78.

[4] K. Takada, H. Kajiwara, N. Imamura, *J. Nat. Prod.* **2010**, *73*, 698–701.

[5] L. Ding, J. Münch, H. Goerls, A. Maier, H. H. Fiebig, W. H. Lin, C. Hertweck, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6685–6687.

[6] L. Ding, A. Maier, H. H. Fiebig, W. H. Lin, C. Hertweck, *Org. Biomol. Chem.* **2011**, *9*, 4029–4031.

[7] L. Ding, A. Maier, H. H. Fiebig, H. Görls, W. H. Lin, G. Peschel, C. Hertweck, *Angew. Chem.* **2011**, *123*, 1668–1672; *Angew. Chem. Int. Ed.* **2011**, *50*, 1630–1634.

[8] a) W. N. Hunter, *J. Biol. Chem.* **2007**, *282*, 21573–21577; b) J. S. Dickschat, *Nat. Prod. Rep.* **2011**, *28*, 1917–1936.

[9] K. Jakobi, C. Hertweck, *J. Am. Chem. Soc.* **2004**, *126*, 2298–2299.

[10] B. Gust, G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton, K. F. Chater, *Adv. Appl. Microbiol.* **2004**, *54*, 107–128.

[11] A. Schenk, Z. Xu, C. Pfeiffer, C. Steinbeck, C. Hertweck, *Angew. Chem.* **2007**, *119*, 7165–7168; *Angew. Chem. Int. Ed.* **2007**, *46*, 7035–7038.

[12] Z. A. Hughes-Thomas, C. B. Stark, I. U. Böhm, J. Staunton, P. F. Leadlay, *Angew. Chem.* **2003**, *115*, 4613–4616; *Angew. Chem. Int. Ed.* **2003**, *42*, 4475–4478.

[13] B. M. Harvey, T. Mironenko, Y. Sun, H. Hong, Z. Deng, P. F. Leadlay, K. J. Weissman, S. F. Haydock, *Chem. Biol.* **2007**, *14*, 703–714.

[14] B. Julien, Z. Q. Tian, R. Reid, C. D. Reeves, *Chem. Biol.* **2006**, *13*, 1277–1286.

[15] F. Forouhar, J. L. Anderson, C. G. Mowat, S. M. Vorobiev, A. Hussain, M. Abashidze, C. Bruckmann, S. J. Thackray, J. Seetharaman, T. Tucker, R. Xiao, L. C. Ma, L. Zhao, T. B. Acton, G. T. Montelione, S. K. Chapman, L. Tong, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 473–478.

[16] R. Uchida, H. Tomoda, S. Omura, *J. Antibiot.* **2006**, *59*, 298–302.

[17] S. Saikia, E. J. Parker, A. Koulman, B. Scott, *FEBS Lett.* **2006**, *580*, 1625–1630.

[18] T. Itoh, K. Tokunaga, Y. Matsuda, I. Fujii, I. Abe, Y. Ebizuka, T. Kushiro, *Nat. Chem.* **2010**, *2*, 858–864.

[19] T. Dairi, *J. Antibiot.* **2005**, *58*, 227–243.

[20] Y. Liu, W. W. J. McWhorter, C. E. Hadden, *Org. Lett.* **2003**, *5*, 333–335.

[21] B. D. Ensley, B. J. Ratzkin, T. D. Osslund, M. J. Simon, L. P. Wackett, D. T. Gibson, *Science* **1983**, *222*, 167–169.

[22] K. Y. Choi, D. Kim, S. C. Koh, J. S. So, J. S. Kim, E. Kim, *J. Microbiol.* **2004**, *42*, 160–162.

[23] S. Kosono, M. Maeda, F. Fuji, H. Arai, T. Kudo, *Appl. Environ. Microbiol.* **1997**, *63*, 3282–3285.

[24] a) M. Daum, S. Herrmann, B. Wilkinson, A. Bechthold, *Curr. Opin. Chem. Biol.* **2009**, *13*, 180–188; b) D. E. Cane, H. Ikeda, *Acc. Chem. Res.* **2012**, *45*, 463–572; c) C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall, J. S. Dickschat, *ChemBioChem* **2012**, *13*, 202–214.