Alkaloid Biosynthesis

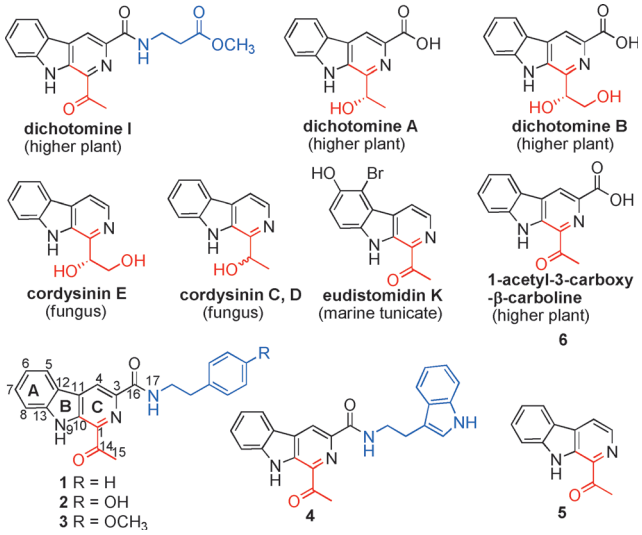
DOI: 10.1002/anie.201303449

Discovery of McbB, an Enzyme Catalyzing the β-Carboline Skeleton Construction in the Marinacarboline Biosynthetic Pathway**

*Qi Chen, Changtao Ji, Yongxiang Song, Hongbo Huang, Junying Ma, Xinpeng Tian, and Jianhua Ju**

The β-carboline (βCs) alkaloids possess a tricyclic pyrido[3, 4-*b*]indole ring system and are widely distributed in nature; the heterocyclic skeleton endows the βCs with antiallergic, antiviral, anti-inflammatory, antibacterial, and antitumor activities, and enables interactions with various receptors leading to neurotoxin and neuroprotectant activities.[1] However, only a few types of enzymes are known to catalyze the Pictet–Spengler (PS) reaction, enabling the biosynthesis of tetrahydro-βC scaffolds. Those characterized enzymes include the strictosidine synthases from various higher plants that are known to catalyze PS condensation of tryptamine and the monoterpenoid aldehyde secologanin to form 3α(*S*)-strictosidine, which serves as the precursor for more than 2500 indole alkaloids.[2] The C-domain of the non-ribosomal peptide synthetase SfmC from *Streptomyces lavendulae* also is known to mediate PS cyclization during the course of saframycin biosynthesis.[3]

There exists a small group of C-ring unsaturated βCs in which C1 is substituted by a two-carbon unit found in various organisms, including medicinal plants,[4] marine tunicate,[5] fungi,[6] and actinomycetes (Scheme 1).[7] Such compounds have attracted significant interest from synthetic chemists.[8] However, the enzymes driving biosynthesis of this group of molecules have not yet been deciphered. We have isolated marinacarbolines A–D (MCBs, **1**–**4**) and 1-acetyl-β-carboline (**5**) from the deep South China Sea-derived *Marinactinospora thermotolerans* SCSIO 00652,[9] thus providing us an excellent opportunity to explore the biosynthetic machinery of this interesting scaffold at a genetic level. Notably, compounds **1**–**4** are all associated with activities against *Plasmodium falciparum* 3D7, a drug-sensitive line, and Dd2, a multi-drug-resistant line.[9] Herein, we report the identification of a gene cluster consisting of three genes *mcbABC* responsible for MCB scaffold biosynthesis, and characterization of McbB, a novel enzyme mediating β-carboline core construction involving a PS cyclization/decarboxylation/oxidation process.



Scheme 1. Structures of marinacarbolines A–D (**1**–**4**) and structurally related β-carboline alkaloids from higher plants, marine sponges, fungi, and actinomycetes.

Considering that the βC metabolites **1**–**4** all contain an aromatic substituted ethylamine group and that the C-ring is unsaturated, we initially hypothesized that a decarboxylase and a dehydrogenase/oxidase are required for MCB scaffold construction. Previously, we shotgun sequenced the whole genome of *M. thermotolerans* SCSIO 00652 and identified the biosynthetic gene clusters governing the biosynthesis of nucleoside antibiotic A201A and methylpendolmycin/pendolmycin.[10] Bioinformatics analysis of the genome revealed eight DNA segments distributed on scaffolds 1, 2, 11, 20, 21, 27, 47, and 93 coding for both types of enzymes, and therefore having a potential role in MCB biosynthesis. Further analysis revealed that the DNA segment on scaffold 27 harbors a fatty acid CoA ligase (Supporting Information, Table S1), potentially involved in amide bond formation on the way to **1**. To explore the possible involvement of these genes in MCB biosynthesis, a pOJ436 vector-based genomic library was constructed and cosmids containing the DNA segment harboring the three types of enzymes were screened. Two positive cosmids, 1610B and 273H, were transformed into *Streptomyces lividans* TK64 through direct conjugation for gene cluster expression. HPLC analyses of the fermentation extracts revealed that all the exconjugants harboring either

[*] Q. Chen, C. Ji, Dr. Y. Song, Dr. H. Huang, Dr. J. Ma, Dr. X. Tian, Prof. Dr. J. Ju
CAS Key Laboratory of Marine Bio-resources Sustainable Utilization, Guangdong Key Laboratory of Marine Materia Medica, RNAM Center for Marine Microbiology, South China Sea Institute of Oceanology, Chinese Academy of Sciences
164 West Xingang Road, Guangzhou 510301 (China)
E-mail: jju@scsio.ac.cn
Q. Chen, C. Ji
University of Chinese Academy of Sciences
19 Yuquan Road, Beijing 110039 (China)

[**] This work was financially supported by grants from MOST (2012AA092104, 2010CB833805), NSFC (31290233, 31000051, 41106138), and CAS (KSCX2-EW-G-12, SQ201015).
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201303449.

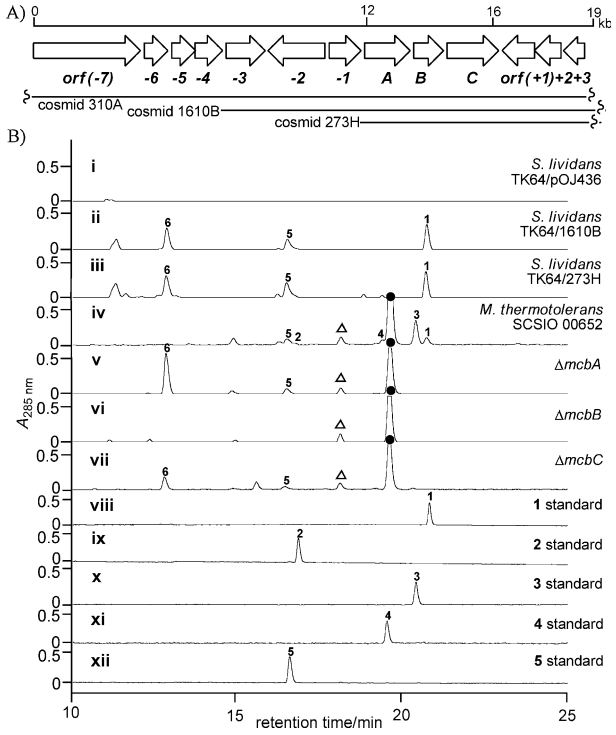


Figure 1. A) Organization of the marinacarboline gene cluster in *M. thermotolerans* SCSIO 00652. B) Metabolite profiles upon HPLC analysis using an Alltima C18 column: i)–iii) *S. lividans* TK64 harboring pOJ436 vector, cosmids 1610B and 273H, respectively; iv)–vii) *M. thermotolerans* SCSIO 00652 wild-type and its mutant strains; viii)–xii) authentic standards of **1**–**5**. ● methylpendolmycin, △ pendolmycin.[10b]

cosmid successfully produce **1** and **5**, and another analogue **6**, whereas the control *S. lividans* TK64 with an empty pOJ436 vector was unable to produce these compounds (Figure 1 B, traces i–iii). These findings suggest that cosmids 1610B and 273H harbor the complete gene set responsible for MCB biosynthesis. Indeed, end sequencing, along with bioinformatics analysis, indicated that a 11.2 kb DNA sequence consisting of 13 open reading frames (orfs) on scaffold 27 most likely contain all the genes accounting for MCB biosynthesis (Supporting Information, Table S1; *orf(−4)* through *orf(−7)* not shown). The nucleotide sequences have been deposited in the GenBank with accession number KC541560.

To precisely investigate how many genes are involved in MCB biosynthesis, we first determined the boundaries of the MCB gene cluster. We screened the previously constructed DNA library derived from supercos1 vector[10] and used the positive cosmid 310A to inactivate the 13 genes by replacing them individually with an apramycin resistance gene cassette using λ-RED-mediated PCR-targeting mutagenesis methodology.[10,11] The four-gene cassette *orf(−7)*, *orf(−6)*, *orf(−5)* and *orf(−4)*, encoding nitrate reductase α, β, γ and δ subunits respectively, was proposed not to be involved in MCB biosynthesis; hence these four genes were deleted in combination. The corresponding mutant strain is designated *orf((−7)-(−4))*. The resulting 10 mutant strains were selected on the basis of kanamycin^S and apramycin^R phenotypes; genotypes were subsequently confirmed by PCR analysis (Supporting Information, Figures S1–S10). All mutant strains were fermented alongside the wild-type control and the resulting metabolites analyzed by HPLC. The upstream cluster boundary was determined by gene inactivations of *orf((−7)-(−4))*, *orf(−3)*, *orf(−2)*, and *orf(−1)*; none of the inactivated mutants accumulated MCBs or related congeners at rates or efficiencies significantly different from those of the wild-type strain. Similarly, the downstream cluster boundary was determined by disruption of *orf(+1)*, *orf(+2)* and *orf(+3)*; none of these mutants displayed changes in MCB production relative to wild-type (Supporting Information, Figure S11). Thus, the regulatory, resistance, and oxidoreductase coding genes (two of each category) listed in the Supporting Information, Table S1 seem not to play a role in MCB biosynthesis in the native *M. thermotelerans* SCSIO 00652 producer or their functions are supplemented by other corresponding genes in the genome. In contrast, the remaining three genes, *mcbABC*, appear to be solely responsible for MCB scaffold construction.

McbA exhibits 42 % identity and 55 % similarity with SRIM_40803, a putative fatty acid CoA ligase from *Streptomyces rimosus subsp. rimosus* ATCC 10970. McbB exhibits 48 % identity and 62 % similarity with BBA_06474, an enzyme with unknown biosynthetic function from *Beauveria bassiana* ARSEF 2860. McbC displays 32 % identity and 49 % similarity with Gad2, an assigned glutamate decarboxylase from *Beggiatoa* sp. PS. Generation of Δ*mcbB* mutant and subsequent metabolite profiling revealed that *mcbB* inactivation abolishes production of **1**–**5** and related analogues demonstrating that McbB is absolutely essential for β-carboline core construction (Figure 1 B, trace vi). The Δ*mcbA* mutant also failed to produce MCBs **1**–**4** but accumulated compound **6** as its major product, along with a minor product **5** (Figure 1 B, trace v). Metabolite profiling of the Δ*mcbC* mutant afforded an HPLC profile similar to that observed when evaluating fermentations of the Δ*mcbA* mutant although the yields of **5** and **6** were relatively low in the Δ*mcbC* case (Figure 1 B, trace vii). Compound **6** has a molecular formula of $C_{14}H_{10}N_2O_3$ as determined by HRMS and was identified as 1-acetyl-3-carboxy-β-carboline, following MS, ^{1}H and ^{13}C NMR spectroscopy, and HMBC data analyses (Supporting Information). Importantly, **6** has been previously reported to be isolated from higher plants[4b,c] and compound **5** was identified as 1-acetyl-β-carboline by comparison with an authentic sample. Consequently, in vivo gene inactivation results suggest that McbB plays a vital role in βC skeleton assembly and that the effects of McbA and McbC knockout are consistent with inactivation of CoA ligase and/or decarboxylase (of Phe/Tyr/Trp) functions. Ultimately, assignments of McbA as a CoA ligase and McbC as a decarboxylase were achieved on the basis of bioinformatics.

To validate that *mcbA*, *mcbB*, and *mcbC* are sufficient for MCB scaffold biosynthesis, we next carried out heterologous expression (either in combinations or individually) of *mcbABC*, *mcbAB* and *mcbB* in *E. coli* BL21. The genes *mcbABC*, *mcbAB* and *mcbB* were cloned into *Nde*I and *Hin*dIII sites of the pET28a(+) vector to generate pET28a(+)/*mcbABC*, pET28a(+)/*mcbAB* and pET28a(+)/

mcbB; these were then transformed into *E. coli* BL21(DE3), respectively. The three strains were cultured in LB medium to an OD_{600} of 0.6, then induced with 0.05 mM IPTG and cultured for another 12 h at 28 °C. HPLC analyses of the culture extracts revealed that: 1) MCBs **1** and **4** were produced as the major product in *E. coli* BL21/pET28a (+)/*mcbABC*; 2) MCBs **1**, **2**, and **4** were produced in *E. coli* BL21/pET28a (+)/*mcbAB*; and 3) the major product **6** and two minor product **5** and **8** were produced in *E. coli* BL21/pET28a (+)/*mcbB* (Figure 2, traces ii–iv). Scaled-up fermentation of *E. coli* BL21/pET28a (+)/*mcbB* (4.5 L) and subsequent isolation afforded analytically pure quantities of **5**, **6**, and **8**. HRMS of **8** yielded a molecular formula of $C_{13}H_{10}N_2O_2$, 28 mass units smaller than that of **6**. The ^1H, ^{13}C, and HMBC NMR data analyses of **8** (Supporting Information) led us to assign its structure as 1-acetyl-3-hydroxy-β-carboline (Scheme 1). These results support the in vivo inactivation results and further demonstrate that the minimal three-gene cassette *mcbABC* is sufficient for MCB scaffold biosynthesis (**1**, **2** and **4**), and that McbB alone is able to construct the βC skeleton intermediate **6**, functioning as a Pictet–Spenglerase that executes C-ring closure and desaturation but that also leaves open the question of whether or not the C_3 unit (C1, C14, and C15) is a biosynthetic precursor to **6**.

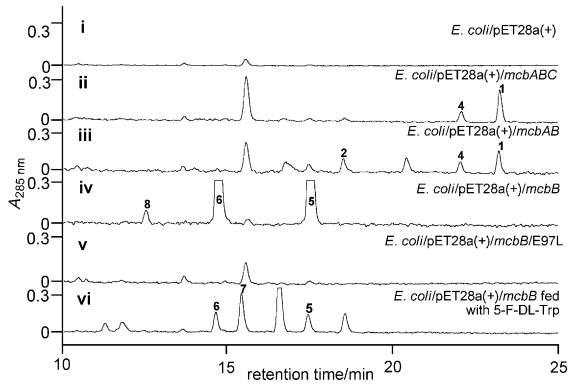
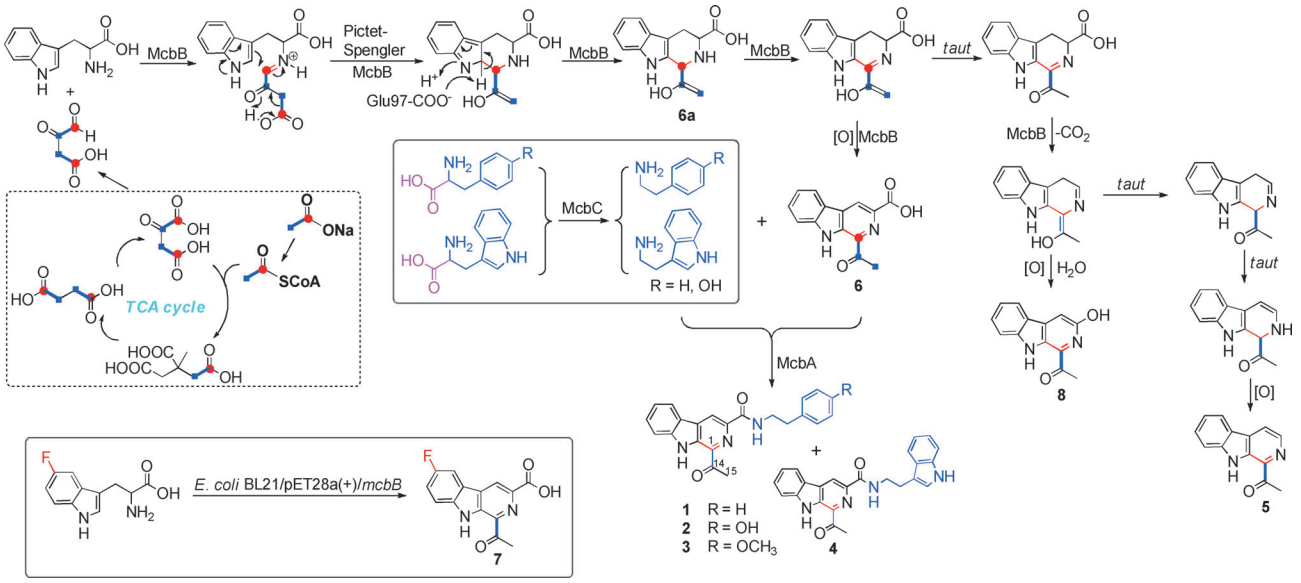


Figure 2. HPLC profiles of culture extracts. i)–v) Recombinant *E. coli* BL21 strains harboring various plasmids, vi) *E. coli* BL21/pET28a (+)/*mcbB* fed with 5-F-DL-Trp.

To explore the biosynthetic precursors that McbB may utilize to biosynthesize βC core **6**, we then performed feeding experiments with 5-fluoro-Trp and ^{13}C-labeled precursors. *E. coli* BL21/pET28a(+)/*mcbB* was cultured, induced with IPTG, and fed with 5-fluoro-DL-Trp. The culture extract, upon HPLC analysis (Figure 2, trace vi), clearly yielded a new product **7** that shows UV absorption bands similar to those of **6** and that has a molecular weight of 272.1 corresponding to 6-fluoro substituted **6**. Scaled-up fermentation (1.5 L) and subsequent HPLC-guided purification led to the isolation of analytically pure **7** for structure elucidation. HRMS, ^1H, ^{13}C, and HMBC NMR data analyses permitted unambiguous assignment of **7** as 1-acetyl-3-carboxy-6-fluoro-β-carboline (Supporting Information). We initially envisioned that the remaining three-carbon unit (C1, C14, and C15) of **6** might originate from a direct C_3 precursor such as pyruvate (from glycolysis). To test this hypothesis, [U-$^{13}C_6$] glucose was fed to the *E. coli* BL21/pET28a (+)/*mcbB* strain and the resultant product **6** was purified. Measurement of the ^{13}C NMR spectrum of **6** revealed that all of the C_3 unit (C1, C14, and C15) was isotopically enriched. However, subsequent 2D INADEQUATE experiment of **6** disclosed that direct correlations were only observed between C1 and C14, indicating that these two carbons are derived from a C_2 unit; the C15 methyl group did not show an apprent correlation with C14. Consequently, feeding experiments with ^{13}C-labeled acetate to *E. coli* BL21/pET28a (+)/*mcbB* were conducted and the corresponding product **6** was purified. Inspection of the ^{13}C NMR spectrum (Supporting Information) of corresponding product **6** revealed that: 1) feeding with [1-^{13}C] acetate led to C1 enrichment in **6**; 2) feeding with [2-^{13}C] acetate led to C14 and C15 enrichments in **6**; and 3) feeding with [1, 2-^{13}C] acetate led to C1, C14, and C15 enrichments in **6**; C1 and C14 appeared as coupled doublets ($^1J_{CC} = 63.1$ Hz) whereas C15 appeared as a singlet. These feeding experiments collectively demonstrate that: 1) the βC core **6** originates from Trp and two acetate units; 2) the C1 and C14 originate from an intact acetate unit; and 3) C15 originates from C2 of the acetate. The labeling pattern of **6** strongly indicate that oxaloacetaldehyde derived from oxaloacetic acid which originates from TCA cycle[12] might be used as the direct precursor, and the single McbB-catalyzed process on the way to **6** involves a Pictet–Spengler cyclization, a decarboxylation reaction, and C-ring oxidation, as shown in Scheme 2.

Finally, we generated point mutations in McbB to probe which amino acid residues are important for catalytic function by expressing the mutated McbB proteins in *E. coli* BL21 and then assaying for production of **6**. BLAST analyses revealed that McbB shows high identity and similarity to a group of proteins primarily annotated from fungi and bacteria whole genome sequences; no conserved domains could be found from this class of uncharacterized proteins. We aligned McbB with ten closely related proteins (Supporting Information, Figure S12) and analyzed the conserved amino acid residues. Site-directed mutagenesis of 42 highly or moderately conserved amino acid residues in McbB was carried out using the QuickChange site-directed mutagenesis kit or overlap extension-PCR methods.[13] Mutant plasmids were transformed into *E. coli* BL21 (DE3) and the resulting cells fermented with IPTG induction; wild-type McbB-bearing cells were processed in an identical fashion thus serving as an enzymatically intact control. Fermentation broths were extracted with EtOAc and the resulting extracts then analyzed by HPLC. These HPLC results revealed that the E97L mutation in McbB completely abolished generation of **6** and minor metabolite **5** (Figure 2, trace v). The yields of these two compounds were sharply decreased to 1–2% in the N24L, L27V, Q42I, H44F, S48A, S83A, Q86I, H87F, K91A, L115A, E181L, and R212L mutants; compound production by the I6V, L36A, D71L, G195A, Y53F, Y38F, Y92F, E110L, and E170L mutants was diminished to 10–40% of that found with the wild-type producer. All the remaining 20 mutants failed to show obvious changes in metabolite yields relative to wild-type (Supporting Information, Figure S14). Consequently, the N-terminus region of McbB appears to contain more residues



Scheme 2. Marinacarbolines **1**–**4** are biosynthesized by McbA, McbB, and McbC, and proposed McbB-catalyzed biosynthetic pathway for **6** using Trp and acetate as precursors. Labeling pattern of **6** derived from the feeding of ^{13}C-labeled sodium acetate is shown.

involved in McbB bioactivity; the conserved amino acid residue E_{97} is absolutely required and may serve as the catalytic center of McbB.

In light of the above data, the chemistry carried out by McbB appears to involve a domino reaction sequence composed of initial Pictet–Spenger condensation followed by decarboxylation, and a C-ring desaturation reaction on the way to major compound **6** and minor metabolites **5** and **8**. We envision the chemistry of MCB scaffold construction to proceed as depicted in Scheme 2. Oxaloacetaldehyde, derived from oxaloacetic acid, might be utilized as the direct precursor for initial Schiff base formation, as per standard Pictet–Spengler chemistry. Concomitant decarboxylation leads and C-ring installation affords key intermediate **6a**. A related means of effecting C–C bond cleavage has been recently implicated in the biosynthesis of aziridine-containing natural products azicemicin A and B.[12] Importantly, **6a** can proceed to compounds **6**, **8**, and **5** through sequences that include: 1) C-ring aromatization via two-fold oxidation (affording **6**); 2) C-ring oxidation followed by decarboxylation, hydration, and ultimate C-ring aromatization (affording **8**); and 3) C-ring oxidation followed by decarboxylation and C-ring aromatization (affording **5**). It is important to note that throughout the pathway to these products, oxidative transformations play a central role, some of which may proceed spontaneously, as evidenced by chemical synthesis of **5** and **8** using L-Trp, methylglyoxal, and *p*-toluenesulfonic acid (1.0 equiv) as a catalyst in yields of 5% and 13%, respectively.[8b] The alkaloids harman and eleagnine are other examples of natural products whose β-carboline scaffold is attributed to oxidative transformations of indole intermediates.[14] Likewise, we envision that generation of the present MCB scaffold relies, in large measure, on the oxidative amenability of indole **6a** and related compounds. No doubt aided by this reactivity, it is remarkable to discover the diversity of chemistries enabled by the multifaceted McbB protein on the way to βC scaffold **6**; the catalytic mechanisms driven by McbB are the subject of future exploration.

In conclusion, genome scanning, bioinformatics analysis, gene inactivation, and heterologous expression in *Streptomyces lividans* TK64 and *Escherichia coli* BL21 have enabled us to identify and elucidate three genes, *mcbABC*, that drive the biosynthesis of MCBs. When expressed on its own in *E. coli*, *mcbB*, produced 1-acetyl-3-carboxy-β-carboline (**6**) as the major product and two other minor products 1-acetyl-β-carboline (**5**) and 1-acetyl-3-hydroxy-β-carboline (**8**), thus highlighting it as a novel enzyme for β-carboline scaffold construction. Feeding experiments with 5-F-Trp and ^{13}C-labeled glucose and acetate reveal that **6** is derived from Trp and two equivalents of acetate as precursors. Generation of **6** by McbB involves a Pictet-Spengler cyclization, a decarboxylation and C-ring oxidation process. Site-directed mutagenesis experiments with McbB reveal that E_{97} is absolutely required for biochemical activity. All in all, these studies enhance our understanding of the genetic and biochemical aspects of MCB construction thereby expanding our repertoire of methods by which to search for and to develop new βC-containing natural products.

Received: April 23, 2013
Revised: June 9, 2013
Published online: August 1, 2013

Keywords: biosynthesis · β-carbolines · marinacarbolines · natural products · Pictet–Spengler reaction

[1] a) R. Cao, W. Peng, Z. Wang, A. Xu, *Curr. Med. Chem.* **2007**, *14*, 479–500; b) T. Herraiz in *Isoquinolines and beta-carbolines as neurotoxins and neuroprotectants* (Eds.: L. Antkiewicz-Michaluk, H. Rommelspacher), Springer, New York, **2012**, pp. 77–103.
[2] a) J. Stöckigt, A. P. Antonchick, F. Wu, H. Waldmann, *Angew. Chem.* **2011**, *123*, 8692–8719; *Angew. Chem. Int. Ed.* **2011**, *50*,

8538 – 8564; b) L. F. Szabó, *Molecules* **2008**, *13*, 1875 – 1896; c) J. J. Maresh, L. A. Giddings, A. Friedrich, E. A. Loris, S. Panjikar, B. L. Trout, J. Stökigt, B. Peters, S. E. O'Connor, *J. Am. Chem. Soc.* **2008**, *130*, 710 – 723; d) P. Bernhardt, E. McCoy, S. E. O'Connor, *Chem. Biol.* **2007**, *14*, 888 – 897; e) X. Y. Ma, S. Panjikar, J. Koepke, E. Loris, J. Stöckigt, *Plant Cell* **2006**, *18*, 907 – 920.

[3] K. Koketsu, K. Watanabe, H. Suda, H. Oguri, H. Oikawa, *Nat. Chem. Biol.* **2010**, *6*, 408 – 410.

[4] a) B. Sun, T. Morikawa, H. Matsuda, S. Tewtrakul, L.-J. Wu, S. Harima, M. Yoshikawa, *J. Nat. Prod.* **2004**, *67*, 1464 – 1469; b) L.-H. Cao, W. Zhang, J.-G. Luo, L.-Y. Kong, *Helv. Chim. Acta* **2012**, *95*, 1018 – 1025; c) F. Faini, R. Torres, F. D. Monache, G. B. Marini-Bettolo, M. Castillo, *Planta Med.* **1980**, *38*, 128 – 132.

[5] T. Suzuki, T. Kubota, J. Kobayashi, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4220 – 4223.

[6] M.-L. Yang, P.-C. Kuo, T.-L. Hwang, T.-S. Wu, *J. Nat. Prod.* **2011**, *74*, 1996 – 2000.

[7] H. J. Shin, H.-S. Lee, D.-S. Lee, *J. Microbiol. Biotechnol.* **2010**, *20*, 501 – 505.

[8] a) A. L. Pumphrey, H. Dong, T. G. Driver, *Angew. Chem.* **2012**, *124*, 6022 – 6025; *Angew. Chem. Int. Ed.* **2012**, *51*, 5920 – 5923; b) M.-L. Yang, P.-C. Kuo, A. G. Damu, R.-J. Chang, W.-F. Chiou, T.-S. Wu, *Tetrahedron* **2006**, *62*, 10900 – 10906; c) J. Baiget, S. Llona-Minguez, S. Lang, S. P. Mackay, C. J. Suckling, O. B. Sutcliffe, *Beilstein J. Org. Chem.* **2011**, *7*, 1407 – 1411.

[9] H. Huang, Y. Yao, Z. He, T. Yang, J. Ma, X. Tian, Y. Li, C. Huang, X. Chen, W. Li, S. Zhang, C. Zhang, J. Ju, *J. Nat. Prod.* **2011**, *74*, 2122 – 2127.

[10] a) Q. Zhu, J. Li, J. Ma, M. Luo, B. Wang, H. Huang, X. Tian, W. Li, S. Zhang, C. Zhang, J. Ju, *Antimicrob. Agents Chemother.* **2012**, *56*, 110 – 114; b) J. Ma, D. Zuo, Y. Song, B. Wang, H. Huang, Y. Yao, W. Li, S. Zhang, C. Zhang, J. Ju, *ChemBioChem* **2012**, *13*, 547 – 552.

[11] a) Y. Zhang, F. Buchholz, J. P. Muyrers, A. F. Stewart, *Nat. Genet.* **1998**, *20*, 123 – 128; b) B. Gust, G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1541 – 1546.

[12] Y. Ogasawara, H.-w. Liu, *J. Am. Chem. Soc.* **2009**, *131*, 18066 – 18068.

[13] a) S. Simionatto, S. B. Marchioro, V. Galli, T. D. Luerce, D. D. Hartwig, A. N. Moreira, O. A. Dellagostin, *J. Microbiol. Methods* **2009**, *79*, 101 – 105; b) X. Mo, J. Ma, H. Huang, B. Wang, Y. Song, S. Zhang, C. Zhang, J. Ju, *J. Am. Chem. Soc.* **2012**, *134*, 2844 – 2847.

[14] R. B. Herbert, J. Mann, *J. Chem. Soc. Perkin Trans. 1* **1982**, 1523 – 1525.

