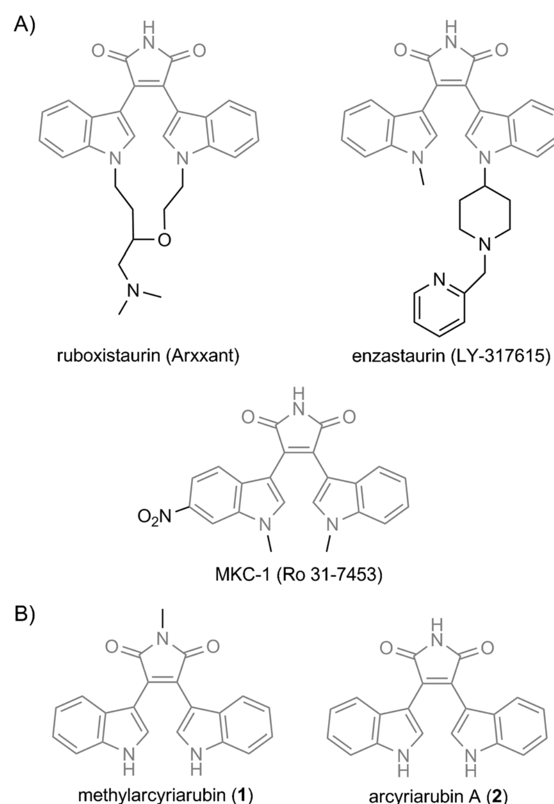
DOI: 10.1002/cbic.201300756

# Characterization of an Environmental DNA-Derived Gene Cluster that Encodes the Bisindolylmaleimide Methylarcyriarubin

Fang-Yuan Chang and Sean F. Brady*[a]

Bisindolylmaleimides represent a naturally occurring class of metabolites that are of interest because of their protein kinase inhibition activity. From a metagenomic library constructed with soil DNA, we identified the four gene *mar* cluster, a bisindolylmaleimide gene cluster that encodes for methylarcyriarubin (**1**) production. Heterologous expression of the *mar* gene cluster in *E. coli* revealed that the Rieske dioxygenase MarC facilitates the oxidative decarboxylation of a chromopyrrolic acid (CPA) intermediate to yield the bisindolylmaleimide core. The characterization of the *mar* cluster defines a new role for CPA in the biosynthesis of structurally diverse bacterial tryptophan dimers.

Bisindolylmaleimide natural products share a common 3,4-di-1*H*-indol-3-yl-1*H*-pyrrole-2,5-dione core structure[1] (Scheme 1, gray). The discovery of the simplest bisindolylmaleimide natural product, arcyriarubin A,[2] from a slime mold (*Arcyria denudate*) spearheaded the extensive study of this class of compounds, with more than 2400 and 4000 bisindolylmaleimide-related references in the PubMed and SciFinder databases, respectively. Bisindolylmaleimide analogues, with activities in cancer,[3] diabetes,[4] and cardiovascular[5] and neurodegenerative[6] disease models have now been synthesized, some of which have advanced into clinical trials[7] (Scheme 1 A). Surprisingly, bisindolylmaleimide biosynthesis has remained uncharacterized.

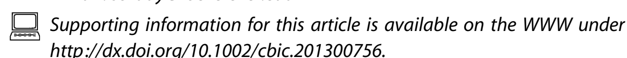



**Scheme 1.** Bisindolylmaleimide compounds, with the 3,4-di-1*H*-indol-3-yl-1*H*-pyrrole-2,5-dione core structure highlighted in gray. A) Synthetic derivatives that have entered into clinical trials for drug therapy. B) Compounds encoded by the eDNA-derived *mar* biosynthetic gene cluster.

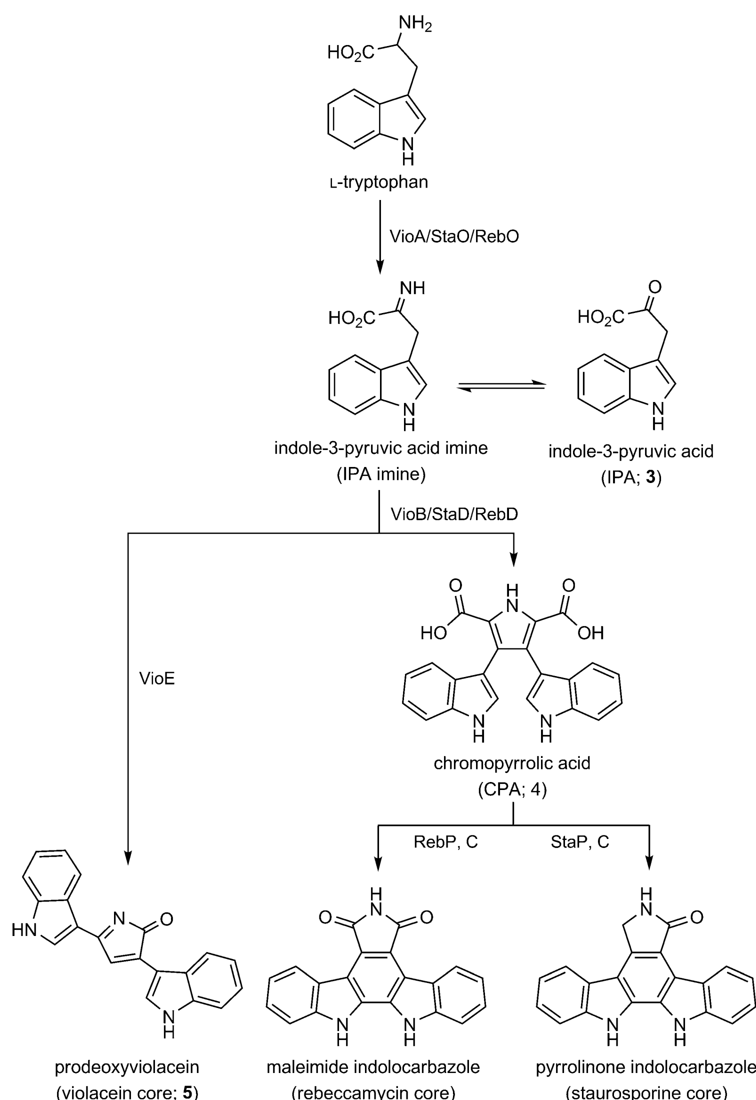
In addition to being isolated from slime molds, bisindolylmaleimides have also now been found in bacterial culture broth extracts, suggesting the presence of bisindolylmaleimide biosynthetic gene clusters in bacterial genomes.[8] Soil is thought to contain thousands of bacterial species per gram, with as much as 99% of these bacteria being recalcitrant to culturing by standard methods.[9] The study of natural microbial communities by culture-independent methods provides a means of systematically exploring large numbers of biosynthetic gene clusters arising from diverse bacterial species. Here, we describe the first functional characterization of bisindolylmaleimide biosynthesis by using a gene cluster recovered from a soil environmental DNA (eDNA) library (Scheme 1 B). Our in vivo analysis of the eDNA-derived methylarcyriarubin (*mar*) gene cluster reveals that a novel dioxygenase, MarC, is responsible for the formation of the bisindolylmaleimide core through the oxidative decarboxylation of a chromopyrrolic acid (CPA) intermediate.

Based on its bisindole chemical structure, bisindolylmaleimide is predicted to arise from the dimerization of two tryptophans. Among structurally characterized bacterial tryptophan dimers, the biosyntheses of violacein and a number of indolocarbazole compounds have been described.[10] In each case, the biosynthesis begins with the oxidation of L-tryptophan by an indole-3-pyruvic acid (IPA) imine synthase (e.g., VioA, RebO, StaO) and then proceeds through dimerization of the IPA imine by a CPA synthase (e.g., VioB, RebD, StaD) to produce CPA (Scheme 2). Because of the structural resemblance of CPA to bisindolylmaleimide, it can be hypothesized that bisindolyl-

[a] F.-Y. Chang, Prof. S. F. Brady
Laboratory of Genetically Encoded Small Molecules
Howard Hughes Medical Institute, The Rockefeller University
1230 York Avenue, New York, NY 10065 (USA)
E-mail: sbrady@rockefeller.edu


Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300756.

 © 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim




**Scheme 2.** Biosynthesis of bacterial tryptophan dimers, namely violacein and indolocarbazole compounds like rebeccamycin and staurosporine, sharing two common initial steps and diverging downstream to yield various chemical structures.

maleimide biosynthesis starts in a similar manner and then diverges to form the bisindolylmaleimide scaffold by oxidative decarboxylation of CPA. Previously reported in vivo and in vitro expression studies on indolocarbazole and violacein biosynthetic genes have found that bisindolylmaleimides (e.g., arcyriarubin A) are neither consumed as substrates nor accumulated as intermediate products in these biosynthetic pathways,[11] suggesting that the diverse collection of known tryptophan dimer modification enzymes is unlikely to be sufficient for producing bisindolylmaleimide structures from CPA.

CPA synthase gene sequences are highly conserved across all known bacterial tryptophan dimer gene clusters. In previous studies, we used degenerate PCR primers designed to recognize conserved regions in CPA synthase genes to screen soil eDNA libraries for potential CPA homologues.[12] In these studies, novel CPA-related amplicons were then used to guide the recovery of eDNA clones containing gene clusters that were bioinformatically predicted to encode a diverse collection of CPA-derived metabolites. Among these is the previously uncharacterized *mar* gene cluster that is encoded on cosmid NM343. Cosmid NM343 was recovered from an eDNA library constructed from Chihuahuan desert soil collected in New Mexico (Figures 1 and S1). The *mar* cluster is predicted to contain four genes: a CPA synthase (*marB*), a dioxygenase (*marC*), a *vioE* homologue (*marE*), and a methyltransferase (*marM*). Dioxygenase MarC is unprecedented in known tryptophan dimer biosynthetic pathways, and we hypothesized that it might direct the formation of a core tryptophan dimer structure that differed from those that had been biosynthetically characterized to date. To investigate this possibility, we chose to study the *mar* cluster by using heterologous expression methods.

Initially, the native *mar* cluster was introduced into a variety of model hosts (e.g., *E. coli*, *Streptomyces* spp., *Burkholderia* spp.) for expression studies, but no clone-specific metabolites were detected in the culture broths. In an effort to address potential transcriptional inefficiencies of *mar* gene cluster promoters in these hosts, we individually cloned the *mar* biosynthetic genes in front of T7 promoters and introduced these constructs into *E. coli*. Unfortunately, this also failed to result in the production of any detectable clone-specific small molecules by *E. coli* cultures (Figure 2, MarBCEM).

Although all of the genetic information required to encode the biosynthesis of a bacterial natural product is generally found clustered on a bacterial chromosome, there are a number of examples in which a required biosynthetic inter-

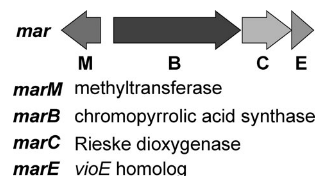



**Figure 1.** eDNA-derived *mar* biosynthetic gene cluster.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

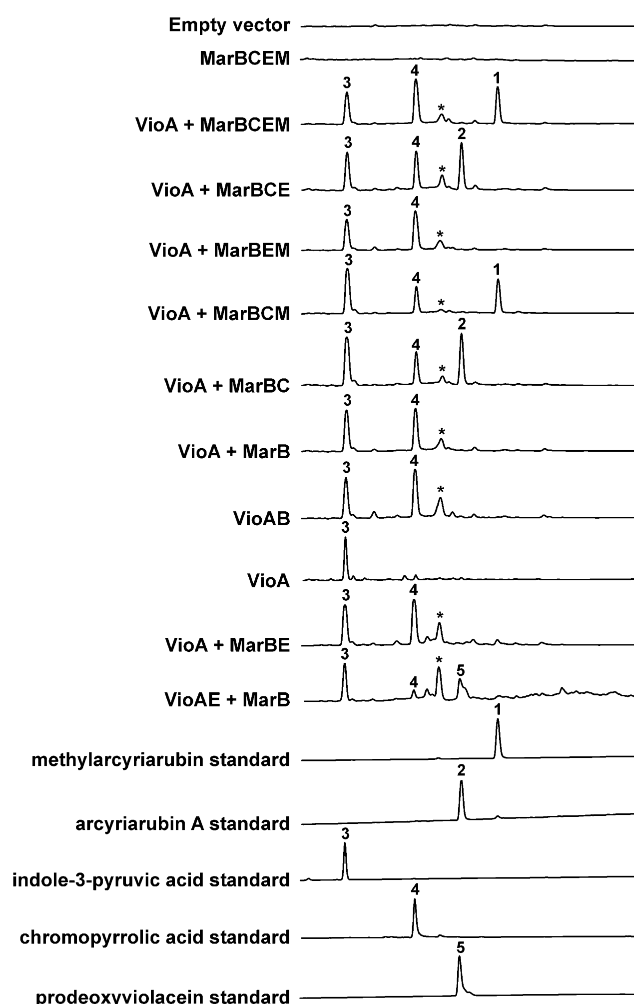



**Figure 2.** HPLC–UV traces of culture broth extracts from *mar* gene cluster expression studies in *E. coli*. IPA imine is reported to undergo spontaneous deamination to form IPA (**3**).[11c,14] The peak marked with an asterisk (*) is an uncharacterized by-product of the coexpression of an IPA imine synthase with a CPA synthase in *E. coli*.

mediate is encoded elsewhere in the genome. For example, genes encoding the biosynthesis of deoxysugar precursors, which are required for many glycosylated compounds, are often found elsewhere in the genome and must, therefore, be introduced *in trans* in heterologous expression studies.[13] With this in mind, we investigated the possibility that the biosynthesis of a precursor required by the *mar* pathway might not be encoded within the *mar* cluster.

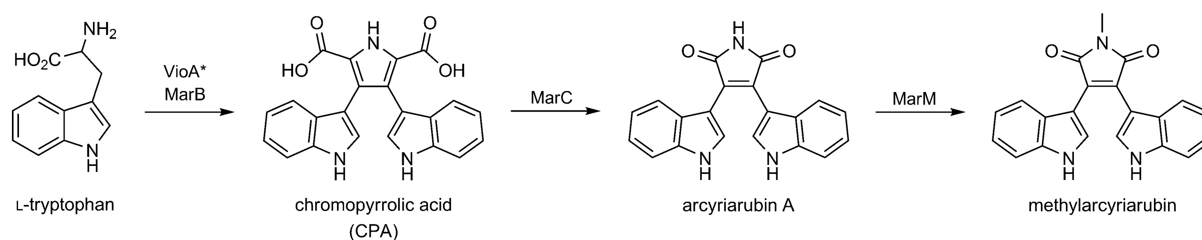
Functionally characterized CPA synthases from other tryptophan dimer biosynthetic clusters have been found to accept oxidized tryptophan (IPA imine), but not tryptophan itself, as a substrate.[14] Neither the *mar* cluster nor the *E. coli* genome contains an IPA imine synthase homologue. Therefore, if MarB functions as a CPA synthase, as predicted by its high sequence identity to known CPA synthases, an IPA imine synthase would have to be supplied *in trans* for *mar* biosynthesis to proceed in a heterologous expression setting. A number of sequenced bacterial genomes contain isolated predicted IPA imine synthase genes, suggesting that IPA imine production might be encoded outside of secondary metabolism in a variety of bacteria. Therefore, we coexpressed the IPA imine synthase *vioA* from the violacein cluster with the rest of the *mar* biosynthetic genes in *E. coli*. This resulted in the production of a clone-specific metabolite (**1**; Figure 2, VioA + MarBCEM), which we had not observed in any previous tryptophan dimer studies, along with the expected tryptophan dimer intermediates IPA (**3**) and CPA (**4**).

Compound **1** (1.6 mg $L^{-1}$) was purified from large-scale cultures of *E. coli* transformed with the VioA + MarBCEM expression constructs. The structure of **1** was solved by using a combination of NMR, UV, and HRMS data (Figures S2, S3) and was determined to be the methylated bisindolylmaleimide methylarcyriarubin (**1**). Although methylarcyriarubin has been made synthetically, to the best of our knowledge, this is the first reported case of methylarcyriarubin being isolated as a natural product.

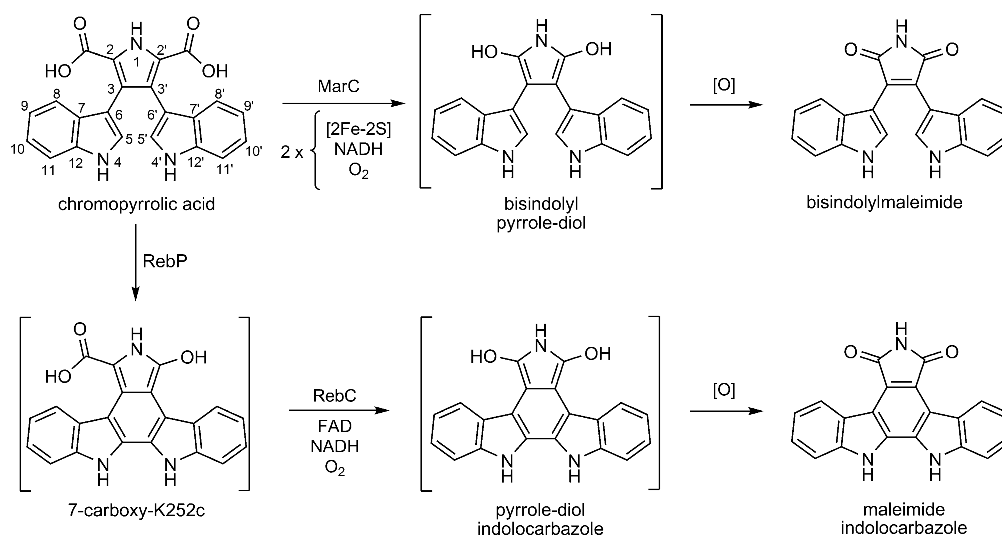
To elucidate the role of the individual *mar* biosynthetic genes in the biogenesis of methylarcyriarubin (**1**), culture broth extracts from *E. coli* strains expressing different combinations of the *mar* biosynthetic genes were characterized in detail. In *E. coli* lacking *marM* (Figure 2, VioA + MarBCE), arcyriarubin A (**2**), the desmethyl form of methylarcyriarubin (**1**), was produced along with **3** and **4**. Without the dioxygenase *marC* (Figure 2, VioA + MarBEM; VioA + MarB), accumulation of **3** and **4** was observed. Replacement of MarB with the characterized CPA synthase from the violacein pathway (VioB) resulted in an identical metabolic profile (Figure 2, VioA + MarB; VioAB). No difference in metabolic profile was observed between *E. coli* expressing all of the *mar* biosynthetic genes (Figure 2, VioA + MarBCEM) and a strain lacking the predicted *vioE* homologue, *marE* (Figure 2, VioA + MarBCM). Based on these expression studies, the biosynthesis of methylarcyriarubin is predicted to share the same two initial biosynthetic transformations as all other biosynthetically characterized bacterial tryptophan dimers (Scheme 3). Specifically, tryptophan is first oxidized to IPA imine by an IPA imine synthase that is found outside of the *mar* cluster in the endogenous host genome. MarB then dimerizes IPA imine to give CPA. The *mar* biosynthetic pathway is then predicted to diverge from the known tryptophan dimer pathway in that MarC appears to function as a bisindolylmaleimide synthase by converting CPA into arcyriarubin A. MarM is predicted to methylate the bisindolylmaleimide to yield methylarcyriarubin (**1**).

MarC is functionally similar to RebC from rebeccamycin biosynthesis in that they both produce a maleimide moiety from CPA. RebC is an FAD-binding monooxygenase that reacts in tandem with RebP to produce maleimide indolocarbazole.[15] Based on cocrystallization studies,[16] the likely substrate of RebC was found to be 7-carboxy-K252c, which is produced by RebP through the C-5 and C-5′ aryl–aryl coupling of CPA. This led to a proposed mechanism for RebC involving hydroxylation of 7-carboxy-K252c at the α-carbon of the carboxyl group to facilitate decarboxylation and yield a pyrrole-diol moiety, followed by oxidative rearrangement to generate the maleimide (Scheme 4). Although MarC is also responsible for generating a maleimide moiety, it is predicted, based on sequence homol-

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim




**Scheme 3.** Biosynthesis of methylarcyriarubin. VioA-like IPA imine synthase (marked with an asterisk) is predicted to be encoded elsewhere in the endogenous host genome.




**Scheme 4.** Comparison of the proposed enzymatic oxidative mechanism between bisindolylmaleimide and indolocarbazole in the biosynthesis of a maleimide moiety.

ogy, to be a Rieske-type dioxygenase. Rieske-type dioxygenases contain a [2Fe–2S] iron–sulfur cluster, instead of FAD, as the cofactor.[17] In MarC-dependent bisindolylmaleimide biosynthesis, we propose that MarC hydroxylates the C2=C3 and C2′=C3′ olefins of CPA in two successive oxidations to facilitate two decarboxylations and generate the pyrrole-diol that can similarly undergo oxidative rearrangement to yield the maleimide, without relying on a second enzyme (Scheme 4). Although our data supports this simple model, the involvement of unknown host factors cannot be ruled out.

Based on our in vivo analyses, no function can yet be assigned to MarE, a predicted VioE homologue. Gel analysis of the soluble protein extract of *marE* harboring *E. coli* confirms that soluble MarE is produced in this system (Figure S4). In violacein biosynthesis, VioE is predicted to produce prodeoxyviolacein (**5**) by interacting with a transient intermediate of unknown structure that is produced by CPA synthase.[18] VioE is a small protein (191 aa) that lacks any functionally characterized homologues, any known catalytic residues, or any recognized cofactor-binding motifs. Accordingly, the mechanistic details of its role in violacein biosynthesis remain unclear. MarE shares high sequence identity to VioE; however, in vivo, MarE cannot complement the function of VioE in the production of prodeoxyviolacein (**5**; Figure 2, VioA+MarBE versus VioAE+MarB). Whether MarE is inactive in the *mar* cluster, plays a role that is redundant in *E. coli*, or functions in the biosynthesis of an as yet unidentified metabolite is subject for further investigation.

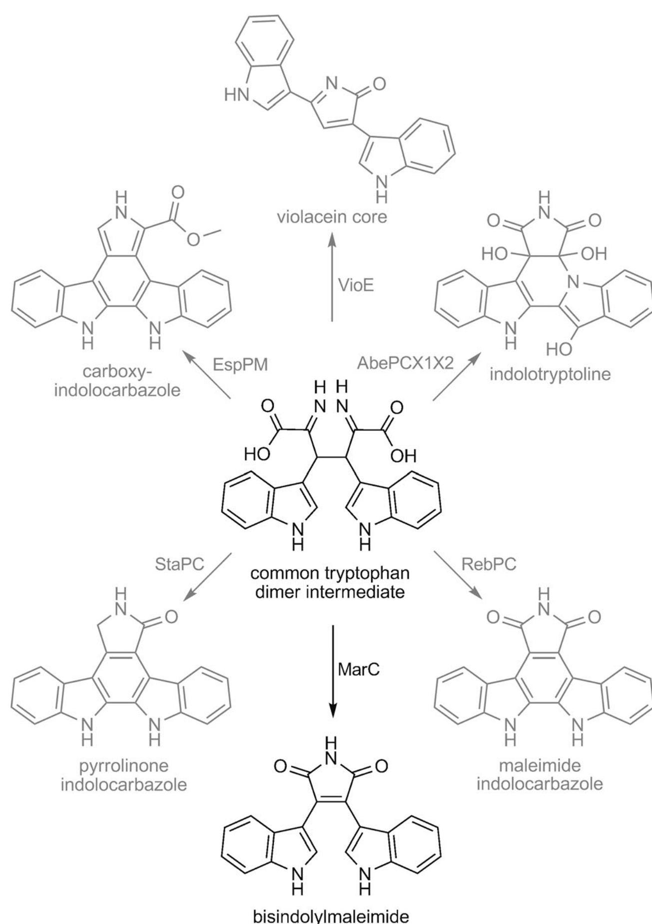
A diverse collection of bisindolylmaleimides have been chemically synthesized and tested for bioactivity, with a particular focus on kinase inhibitory activity.[19] Arcyriarubin A (**2**) is a sub-micromolar inhibitor of protein kinase C ($IC_{50}$: 0.55 μM).[20] Interestingly, the addition of the N-methyl to give methylarcyriarubin (**1**) abolishes protein kinase C inhibitory activity ($IC_{50}$: >100 μM)[20] but leads to activity against mitogen-stimulated protein kinase $p70^{s6k}/p85^{s6k}$ ($IC_{50}$: 8 μM).[21] A number of synthetic bisindolylmaleimide derivatives, including ruboxistaurin, enzastaurin, and MKC-1, have undergone or are currently in clinical trials as potent and specific kinase inhibitors for use as cancer and diabetes therapies.[7]

Indolocarbazole tryptophan dimers, which differ from bisindolylmaleimides by the coupling of the C-5 and C-5′ indole carbons, have also been extensively explored as kinase inhibitors.[22] The additional C–C coupling forces indolocarbazoles to bind in a planar conformation, whereas the more flexible bisindolylmaleimides have been observed to bind in a nonplanar

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

conformation.[23] This conformational flexibility is believed to be responsible for trends observed in tryptophan bioactivities. Bisindolylmaleimides tend to be less potent but more specific kinase inhibitors than their indolocarbazole counterparts.[20,24] Known indolocarbazole biosynthetic enzymes have been used, both in vitro in chemoenzymatic synthesis[25] and in vivo in combinatorial biosynthesis,[26] to generate many unnatural indolocarbazole analogues. Considering the significant clinical relevance of the bisindolylmaleimides, the identification of the *mar* biosynthetic gene cluster should facilitate the generation of additional collections of potentially pharmaceutically relevant tryptophan dimers by biosynthetic approaches.

The dimerization of tryptophan by the tandem action of an IPA imine synthase and a CPA synthase generates a tremendously versatile intermediate (CPA) that appears to serve as a substrate for the biosynthesis of a diverse array of dimer substructures.[27] The characterization of the *mar* cluster adds a new branch to the meta-biosynthetic scheme of bacterial tryptophan dimers (Scheme 5). In this case, the Rieske-type dioxygenase, MarC, is predicted to oxidatively decarboxylate CPA to produce the bisindolylmaleimide core. As only a small subset of known tryptophan dimer natural products have had their biosynthetic gene clusters characterized thus far, it is likely that the current global biosynthetic scheme is still incomplete and that there are more divergent tryptophan dimer pathways awaiting discovery.




**Scheme 5.** Bacterial tryptophan dimer biosynthetic pathways that diverge from a common tryptophan dimer intermediate. Bisindolylmaleimide pathway was found to also branch out via oxidation by MarC.

## Experimental Section

**Soil environmental DNA (eDNA) library construction:** An eDNA library containing more than $1.5\times10^7$ unique clones was constructed from a Chihuahuan Desert (New Mexico; NM library) soil sample as described previously.[28] In brief, the soil was sifted to remove large particulates and heated for 2 h at 70 °C in lysis buffer (100 mM Tris·HCl, 100 mM EDTA, 1.5 M NaCl, 1% (*w/v*) CTAB, 2% (*w/v*) SDS, pH 8.0). The soil particulates were removed by centrifugation (30 min, 4000 *g*, 4 °C). Crude eDNA was precipitated by the addition of isopropanol (0.7 vol) and collected by centrifugation (30 min, 4000 *g*, 4 °C), followed by a wash step with 70% ethanol (10 min, 4000 *g*, 4 °C). The resulting eDNA was resuspended in TE buffer and subjected to agarose gel electrophoresis (1% agarose gel, 16 h, 20 V). Electroelution (2 h, 100 V) of the high molecular weight (HMW; ≥25 kb) compression band yielded purified eDNA. HMW eDNA was concentrated (100 kDa molecular weight cut off), blunt-ended (End-It), ligated into a cosmid vector (pWEB-TNC or pWEB436), packaged into λ phage (MaxPlax), and transfected into *E. coli* (EC100, Epicentre). The transfected cells were inoculated and arrayed into 48-well plates such that each well consisted of ~4–5× $10^3$ clones in LB medium (5 mL) with antibiotic (100 μg mL$^{-1}$ ampicillin or 50 μg mL$^{-1}$ apramycin) and grown overnight. Matching DNA miniprep and glycerol stock pairs were made for each well and arrayed such that sets of eight wells were combined to generate unique library "row pools."

**Screening and recovery of *mar* cluster harboring clone NM343:** The *mar* cluster was found on a cosmid (NM434) recovered from a screen of the New Mexico eDNA library.[12a] In this screen, degenerate PCR primers StaDV-F and StaDV-R (Table 1) were designed to recognize conserved regions in known bacterial CPA synthase genes (accession nos.: *vioB* AF172851.1, *staD* AB088119.1, *rebD* AJ414559.1, *inkD* DQ399653.1, *atmD* DQ297453.1.). These primers were then used to screen the New Mexico eDNA library "row

**Table 1.** PCR primer list.[a]

| Name | Sequence |
|---|---|
| StaDV-F | GTSATGMTSCAGTACCTSTACGC |
| StaDV-R | YTCVAGCTGRTAGYCSGGRTG |
| NM343-F | GAGCAGCTCAAGCTGGTGTG |
| NM343-R | AAGGCCTCGCGAATCTGCTG |
| MarB-F | GAGAccatggcaATGAGCATCCTGGAATTTCCGC |
| MarB-R | GAGAgtcgacCCTCACAAGAGTGGAACGG |
| MarC-F | GAGAtcatgatcATGCTGAGCGCCGAAGACA |
| MarC-R | GAGAaagcttCTCATGCGGTCTCCTTGC |
| MarE-F | GAGAcatATGAGCGCCGCCCGC |
| MarE-R | GAGAggtaccGAGGATTGTTGGTCTGCTGAC |
| MarM-F | GAGAacatgtctATGACAACTCAGGGAACGCC |
| MarM-R | GAGAgtcgacGCTCAGCGTTCTTTCGTGC |
| VioA-F | GAGAcatATGACAAACTATTCCGACATTTGC |
| VioA-R | GAGAcaattgGGAAATCCAGAATGCTCATGC |
| VioB-F | GAGAccatggcaATGAGCATTCTGGATTTCCCC |
| VioB-R | GAGAaagcttTGCATATCAAGCCTCTCTAGAC |
| VioE-F | GCGCcatATGCCGATGCCTGTCCAC |
| VioE-R | GCGCggtaccCACAAACGGAACAGGACTCAGT |

[a] Underlined sequences indicate restriction sites added for cloning purposes.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

pools" for CPA genes. PCR reaction mixtures were as follows: FailSafe PCR Buffer G (10 μL; Epicentre), StaDV-F and StaDV-R primers (0.5 μL each; final concentration of 2.5 μM each), template "row pool" eDNA (0.5 μL; 100 ng), and *Taq* DNA polymerase (0.2 μL; New England Biolabs). PCR cycling conditions: one cycle of 95 °C for 5 min; seven cycles of 95 °C for 30 s, 65 °C for 30 s with 1 °C decrement per cycle to 59 °C, 72 °C for 40 s; 30 cycles of 95 °C for 30 s, 58 °C for 30 s, 72 °C for 40 s; one cycle of 72 °C for 7 min; 4 °C hold. Amplicons of the correct predicted size (~561 base pairs) were gel-purified, reamplified, and sequenced by using the StaDV-F and StaDV-R primers. Those that were confirmed to be CPA synthase gene sequences, based on BLASTX homology search (NCBI), were then used to guide the recovery of the eDNA clone harboring each of the unique CPA sequences identified in the arrayed library. The *mar* cluster-containing clone was recovered from well number 343 of the New Mexico library (NM343) by using a serial dilution approach with primers NM343-F and NM343-R (Table 1).[12a] Clone NM343 was then de novo sequenced at the Sloan Kettering Institute DNA Sequencing Core Facility by using 454 pyrosequencing technology (Roche). The sequence data was assembled by using a GS De Novo Assembler (Roche), annotated by using FGENESB (SoftBerry) for gene prediction and BLASTP (NCBI) for protein homology searches, and deposited in the GenBank database under the accession number KF551863.

**Cloning of genes from the *mar* and violacein (*vio*) gene clusters:** Individual genes were amplified from the *mar* and *vio* clusters by using clones NM343 and CSL51[29] as templates, respectively, by using the Phusion Hot Start Flex DNA polymerase kit (New England Biolabs). Primers are shown in Table 1. PCR cycling conditions: one cycle of 95 °C for 5 min; 30 cycles of 95 °C for 10 s, 62 °C for 30 s, 72 °C for 30 s kb$^{-1}$; one cycle of 72 °C for 7 min; 4 °C hold. Gel-purified amplicons were restriction digested and cloned into the following Duet (Novagen) vectors: MarB, NcoI/SalI sites of pCOLADuet-1; MarC, NcoI/HindIII sites of pETDuet-1; MarE, NdeI/KpnI sites of pETDuet-1; MarM, NcoI/SalI sites of pCDFDuet-1; VioA, NdeI/MfeI sites of pCOLADuet-1; VioB, NcoI/HindIII sites of pCOLADuet-1; VioE, NdeI/KpnI sites of pETDuet-1.

**Heterologous expression of *mar* and *vio* biosynthetic genes:** For expression studies, electrocompetent *E. coli* BL21 cells were transformed with Duet vectors harboring various combinations of *mar* and *vio* biosynthetic genes and grown in LB medium with the required antibiotic combination for selection (30 μg mL$^{-1}$ kanamycin, 100 μg mL$^{-1}$ ampicillin, 100 μg mL$^{-1}$ spectinomycin). Gene expression was induced in cultures grown to an $OD_{600}$ of 0.5 with the addition of IPTG (final concentration of 0.1 mM). After an additional 36 h (200 rpm, 25 °C), the cultures were extracted with ethyl acetate that was acidified to pH ~3–4 with the addition of hydrochloric acid. Extracts dried in vacuo were dissolved in methanol and subjected to reversed phase LC/MS analysis (150×4.6 mm, 5 μm XBridge C18, linear gradient of 80:20 water/methanol to 100% methanol with 0.1% formic acid). Commercially available methylarcyriarubin (i.e., bisindolylmaleimide V, Santa Cruz Biotechnology), arcyriarubin A (i.e., bisindolylmaleimide IV, Santa Cruz Biotechnology), and indole-3-pyruvic acid (Sigma–Aldrich) were run on LC/MS under the same conditions. As CPA and prodeoxyviolacein were not commercially available, these standards were prepared from the heterologous expression of well-defined biosynthetic genes from the violacein pathway. CPA and prodeoxyviolacein were produced from VioAB- and VioABE-expressing *E. coli* cultures, respectively. They were each purified from culture broth extracts by HPLC under conditions based on previous violacein pathway studies.[18] Analytical LC/MS data was acquired on a Micromass ZQ mass spectrometer (Waters).

**Large-scale production and isolation of methylarcyriarubin (1) from *E. coli* cultures expressing *mar* genes:** Cultures of VioA+MarBCEM-expressing *E. coli* BL21 cells (2 L), grown for 36 h (200 rpm, 25 °C) after IPTG induction, were extracted with ethyl acetate (4 L). This extract was initially fractionated by silica gel RediSep flash chromatography (RediSepRf 12 g silica flash column: 3 min 100% chloroform, 27 min linear gradient from 100% chloroform to 85:15 chlorofom/methanol). Compound **1** was eluted with chloroform/methanol (99:1). Compound **1** was then purified (1.6 mg L$^{-1}$) from the 99:1 fraction with water/acetonitrile (65:35) by isocratic reserved-phase HPLC (150×10 mm, 5 μm XBridge C18). An LTQ-Orbitrap mass spectrometer (Thermo Scientific) and a 600 MHz spectrometer (Bruker) were used to acquire HRMS and NMR data, respectively, for structure elucidation studies.

**MarE protein expression analysis:** A liquid culture (100 mL) of *E. coli* harboring *vioE*/pETDuet-1 was grown to an $OD_{600}$ of 0.5. The culture was subsequently split into two; one of which was induced with the addition of IPTG (final concentration of 0.1 mM). Two hours (200 rpm, 37 °C) post-induction, an aliquot (1 mL) was removed from each culture, and the cells were pelleted by centrifugation (1 min, 13000*g*). The cells were resuspended in native purification buffer (200 μL; 0.5 M NaCl, 50 mM $NaH_2PO_4$, pH 8.0) and lysed by sonication (30 cycles of 1 s pulse on and 2 s pulse off, 45% amplitude; Sonic Dismembrator, Fisher Scientific). The cell debris and the insoluble proteins were collected by centrifugation (15 min, 13000*g*, 4 °C). Aliquots (20 μL) of the supernatant were mixed with SDS loading buffer (50 mM Tris·HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02% Bromophenol Blue), heated for 10 min at 95 °C and run on a polyacrylamide gel (4–20% Mini-PROTEAN TGX Gel with Precision Plus Protein Dual Color Standards, Bio-Rad). Gels were stained with Coomassie (Coomassie Brilliant Blue R-250 staining solution, Bio-Rad) and imaged by using Gel Doc XR+ System (Bio-Rad).

## Acknowledgements

*This work was supported by theNational Institutes of Health (GM-77516). S.F.B. is a Howard Hughes Medical Investigator (HHMI) Early Career Scientist.*

**Keywords:** biosynthesis · bisindolylmaleimides · metagenomics · natural products · Rieske dioxygenase

[1] M. Gill, W. Steglich, *Fortschr. Chem. Org. Naturst.* **1987**, *51*, 1–317.
[2] W. Steglich, B. Steffan, L. Kopanski, G. Eckhardt, *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 459–460; *Angew. Chem.* **1980**, *92*, 463–464.
[3] a) J. Zhang, P. L. Yang, N. S. Gray, *Nat. Rev. Cancer* **2009**, *9*, 28–39; b) E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discovery* **2008**, *7*, 608–624; c) B. Pajak, S. Orzechowska, B. Gajkowska, A. Orzechowski, *Adv. Med. Sci.* **2008**, *53*, 21–31.
[4] P. W. Anderson, J. B. McGill, K. R. Tuttle, *Curr. Opin. Nephrol. Hypertens.* **2007**, *16*, 397–402.
[5] R. G. Katare, Z. Zhitian, M. Sodeoka, S. Sasaguri, *Can. J. Physiol. Pharmacol.* **2007**, *85*, 979–985.
[6] R. Asakai, Y. Aoyama, T. Fujimoto, *Neurosci. Res.* **2002**, *44*, 297–304.
[7] a) L. P. Aiello, M. D. Davis, A. Girach, K. A. Kles, R. C. Milton, M. J. Sheetz, L. Vignati, X. E. Zhi, *Ophthalmology* **2006**, *113*, 2221–2230; b) S. Ma, S. T. Rosen, *Curr. Opin. Oncol.* **2007**, *19*, 590–595; c) F. Zagouri, T. N. Sergentanis, D. Chrysikos, M. Filipits, R. Bartsch, *Gynecol. Oncol.* **2012**, *127*, 662–672.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[8] T. Li, Y. Du, Q. Cui, J. Zhang, W. Zhu, K. Hong, W. Li, *Mar. Drugs* **2013**, *11*, 466–488.
[9] V. Torsvik, J. Goksøyr, F. L. Daae, *Appl. Environ. Microbiol.* **1990**, *56*, 782–787.
[10] a) J. M. Pemberton, K. M. Vincent, R. J. Penfold, *Curr. Microbiol.* **1991**, *22*, 355–358; b) H. Onaka, S. Taniguchi, Y. Igarashi, T. Furumai, *J. Antibiot.* **2002**, *55*, 1063–1071; c) C. Sánchez, I. A. Butovich, A. F. Braña, J. Rohr, C. Méndez, J. A. Salas, *Chem. Biol.* **2002**, *9*, 519–531; d) H. T. Chiu, Y. L. Chen, C. Y. Chen, C. Jin, M. N. Lee, Y. C. Lin, *Mol. BioSyst.* **2009**, *5*, 1180–1191; e) Q. Gao, C. Zhang, S. Blanchard, J. S. Thorson, *Chem. Biol.* **2006**, *13*, 733–743.
[11] a) A. R. Howard-Jones, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 12289–12298; b) C. J. Balibar, C. T. Walsh, *Biochemistry* **2006**, *45*, 15444–15457; c) S. Asamizu, S. Hirano, H. Onaka, H. Koshino, Y. Shiro, S. Nagano, *ChemBioChem* **2012**, *13*, 2495–2500.
[12] a) F.-Y. Chang, M. A. Ternei, P. Y. Calle, S. F. Brady, *J. Am. Chem. Soc.* **2013**, *135*, 17906–17912; b) F. Y. Chang, S. F. Brady, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2478–2483; c) F. Y. Chang, S. F. Brady, *J. Am. Chem. Soc.* **2011**, *133*, 9996–9999.
[13] J. A. Salas, C. Méndez, *Trends Microbiol.* **2007**, *15*, 219–232.
[14] A. R. Howard-Jones, C. T. Walsh, *Biochemistry* **2005**, *44*, 15652–15663.
[15] a) A. R. Howard-Jones, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 11016–11017; b) K. Groom, A. Bhattacharya, D. L. Zechel, *ChemBioChem* **2011**, *12*, 396–400; c) S. Asamizu, Y. Shiro, Y. Igarashi, S. Nagano, H. Onaka, *Biosci. Biotechnol. Biochem.* **2011**, *75*, 2184–2193.
[16] a) K. S. Ryan, A. R. Howard-Jones, M. J. Hamill, S. J. Elliott, C. T. Walsh, C. L. Drennan, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15311–15316; b) P. J. Goldman, K. S. Ryan, M. J. Hamill, A. R. Howard-Jones, C. T. Walsh, S. J. Elliott, C. L. Drennan, *Chem. Biol.* **2012**, *19*, 855–865; c) D. P. Ballou, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15587–15588; d) S. C. Tsai, *Chem. Biol.* **2012**, *19*, 787–788.
[17] S. M. Barry, G. L. Challis, *ACS Catal.* **2013**, *3*, 2362–2370.
[18] a) C. Sánchez, A. F. Braña, C. Méndez, J. A. Salas, *ChemBioChem* **2006**, *7*, 1231–1240; b) S. Hirano, S. Asamizu, H. Onaka, Y. Shiro, S. Nagano, *J. Biol. Chem.* **2008**, *283*, 6459–6466; c) K. S. Ryan, C. J. Balibar, K. E. Turo, C. T. Walsh, C. L. Drennan, *J. Biol. Chem.* **2008**, *283*, 6467–6475.
[19] a) B. A. Teicher, *Clin. Cancer Res.* **2006**, *12*, 5336–5345; b) D. Brehmer, K. Godl, B. Zech, J. Wissing, H. Daub, *Mol. Cell. Proteomics* **2004**, *3*, 490–500.
[20] P. D. Davis, C. H. Hill, G. Lawton, J. S. Nixon, S. E. Wilkinson, S. A. Hurst, E. Keech, S. E. Turner, *J. Med. Chem.* **1992**, *35*, 177–184.
[21] N. Marmy-Conus, K. M. Hannan, R. B. Pearson, *FEBS Lett.* **2002**, *519*, 135–140.
[22] H. Nakano, S. Omura, *J. Antibiot.* **2009**, *62*, 17–26.
[23] a) D. Komander, G. S. Kular, A. W. Schüttelkopf, M. Deak, K. R. Prakash, J. Bain, M. Elliott, M. Garrido-Franco, A. P. Kozikowski, D. R. Alessi, D. M. van Aalten, *Structure* **2004**, *12*, 215–226; b) M. Gassel, C. B. Breitenlechner, N. Konig, R. Huber, R. A. Engh, D. Bossemeyer, *J. Biol. Chem.* **2004**, *279*, 23679–23690.
[24] D. Toullec, P. Pianetti, H. Coste, P. Bellevergue, T. Grand-Perret, M. Ajakane, V. Baudet, P. Boissin, E. Boursier, F. Loriolle, L. Duhamel, D. Charon, J. Kirilovsky, *J. Biol. Chem.* **1991**, *266*, 15771–15781.
[25] a) C. S. Zhang, R. L. Weller, J. S. Thorson, S. R. Rajski, *J. Am. Chem. Soc.* **2006**, *128*, 2760–2761; b) C. Zhang, C. Albermann, X. Fu, N. R. Peters, J. D. Chisholm, G. Zhang, E. J. Gilbert, P. G. Wang, D. L. Van Vranken, J. S. Thorson, *ChemBioChem* **2006**, *7*, 795–804.
[26] a) C. Sánchez, L. Zhu, A. F. Braña, A. P. Salas, J. Rohr, C. Méndez, J. A. Salas, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 461–466; b) A. P. Salas, L. Zhu, C. Sánchez, A. F. Braña, J. Rohr, C. Méndez, J. A. Salas, *Mol. Microbiol.* **2005**, *58*, 17–27.
[27] K. Speck, T. Magauer, *Beilstein J. Org. Chem.* **2013**, *9*, 2048–2078.
[28] J. G. Owen, B. V. B. Reddy, M. A. Ternei, Z. Charlop-Powers, P. Y. Calle, J. H. Kim, S. F. Brady, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11797–11802.
[29] S. F. Brady, C. J. Chao, J. Handelsman, J. Clardy, *Org. Lett.* **2001**, *3*, 1981–1984.
[30] a) M. Brenner, H. Rexhausen, B. Steffan, W. Steglich, *Tetrahedron* **1988**, *44*, 2887–2892; b) K. Wang, Z. Liu, *Synth. Commun.* **2010**, *40*, 144–150.

Received: November 30, 2013
Published online on March 19, 2014

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim