

Biosynthetic Characterization and Chemoenzymatic Assembly of the Cryptophycins. Potent Anticancer Agents from *Nostoc* Cyanobionts

Nathan A. Magarvey^{†,¶}, Zachary Q. Beck^{‡,¶}, Trimurtulu Golakoti[§], Yousong Ding[‡], Udo Huber[§], Thomas K. Hemscheidt[§], Dafna Abelson[†], Richard E. Moore[§], and David H. Sherman^{†,*,¶}

[†]Department of Microbiology and BioTechnology Institute, University of Minnesota, Minneapolis–St. Paul, Minnesota 55108, [‡]Life Sciences Institute, Departments of Medicinal Chemistry, Chemistry, and Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109, [§]Department of Chemistry, University of Hawaii, Manoa, Hawaii 96822. [¶]These authors contributed equally to this work.

ABSTRACT The lichen cyanobacterial symbiont *Nostoc* sp. ATCC 53789 and its close relative *Nostoc* sp. GSV 224 are prolific producers of natural products, generating >25 derivatives of the cryptophycin class of secondary metabolites. Cryptophycin 1, the prototypic member of the class, is a potent tubulin-depolymerizing agent, and several semisynthetic derivatives are being developed as anticancer therapeutics. Here we provide a detailed characterization of the cryptophycin metabolic pathway by stable-isotope labeling experiments and through cloning, sequencing, and annotating the cryptophycin biosynthetic gene cluster. A comparative secondary metabolomic analysis based on polyketide (PK)/non-ribosomal peptide gene clusters from the phylogenetically related, non-cryptophycin producing cycad symbiont, *Nostoc punctiforme* ATCC 29133, was used to identify the cryptophycin biosynthetic genes that encompass ~40 kb within the lichen symbiont *Nostoc* sp. ATCC 53789 genome. The pathway encodes a collinear set of enzymes, including three modular PK synthases, two non-ribosomal peptide synthetase modules, and an integrated adenylation/ketoreductase didomain for elaboration of the leucic acid subunit. In addition, genes encoding key tailoring steps, including a FAD-dependent halogenase and CYP450 epoxidase, were identified. The inherent flexibility of the cryptophycin biosynthetic enzymes was harnessed to generate a suite of new analogues by altering the pool of PK starter units and selected amino acid extender groups. Characterization of the cryptophycin CYP450 enabled development of the first stereospecific synthesis of cryptophycin 2, through a tandem chemoenzymatic synthesis from the natural seco-cryptophycin 4 chain elongation intermediate.

Cryptophycins are the largest class of peptolides isolated from cyanobacteria (blue-green algae) to date (1). The lichen cyanobacterial symbiont *Nostoc* sp. ATCC 53789 and its close genetic relative *Nostoc* sp. GSV 224 produce these promising anticancer agents (1–5). Cryptophycin 1 (1), the major representative of >25 naturally occurring analogues, consists of four units, including a phenyl-octenoic acid (unit A) and L-leucic acid (unit D) and two amino acids, 3-chloro-O-methyl-D-tyrosine (unit B) and methyl β -alanine (unit C), linked in a cyclic ABCD sequence (Table 1). All of the other naturally occurring cryptophycins are analogues that differ structurally from 1 by one or two units in the molecule (Table 1). The *Nostoc*-derived cryptophycins exhibit extensive variation, indicating the flexibility and versatility of the biosynthetic pathway (1–5) (Table 1). Natural cryptophycin variants of unit A differ in their oxygenation state (e.g., alkene vs hydroxyl groups and epoxide vs styrene) and double bond configurations (*trans* vs *cis*). Unit B contains phenylalanine variants, unit C includes methyl β -alanine or β -alanine, and unit D involves α -hydroxy acid subunit diversity (Table 1). Another notable cryptophycin variation, which does not commonly stem from other polyketide (PK) and non-ribosomal peptide biosynthetic pathways, is macrocyclic ring size (16- vs 14-membered peptolide rings) (see Table 1) (1, 3, 4).

1 (Table 1), found as the most abundant product from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224, is one of the most potent tubulin destabilizing agents ever discovered (6). It arrests tumor cells at the G₁-M phase, inducing a block in cellular proliferation, and causes hyper-phosphorylation of Bcl-2, triggering the apoptotic cascade (7). Cryptophycins are also attractive as chemotherapeutic agents because they are not substrates for

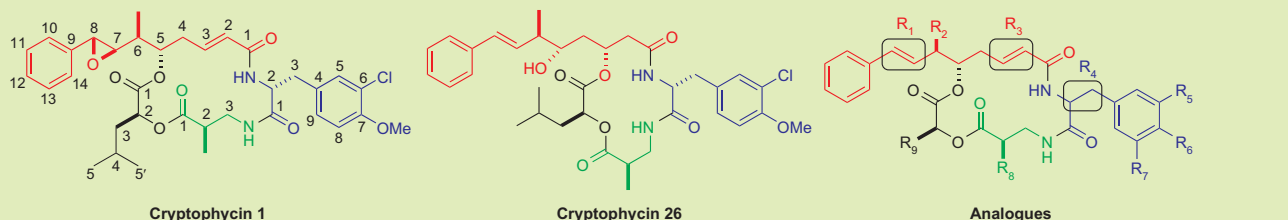
*Corresponding author,
davidhs@umich.edu.

Received for review October 16, 2006
and accepted November 29, 2006.

Published online December 15, 2006

10.1021/cb6004307 CCC: \$33.50

© 2006 by American Chemical Society

TABLE 1. Structural variation of cryptophycins isolated from *Nostoc* strains^a


		Natural cryptophycin analogues																											
		1	2	3	4	16	17	18	19	21	23	24	26	28	29	30	31	38	40	43	45	46	49	50	54	175	176	326	327
R ₁	β-epoxide	•	•			•		•			•	•		•			•							•	•		•	•	•
	<i>trans</i> -Styrene			•	•		•		•	•			•		•	•				•	•	•	•			•			
	α-Epoxide																		•										
R ₂	CH ₃	•	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•		•	•	•	•	•	•	•	•	•	•
	H													•					•										
R ₃	Trans double bond	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
	Cis double bond																												•
	OH											•																	
R ₄	L																					•							
	D	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•	•
R ₅	Cl	•		•		•	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	H		•		•							•								•									
R ₆	OCH ₃	•	•	•	•			•	•	•	•	•	•	•	•	•	•	•	•			•	•	•	•	•		•	•
	OH					•	•													•	•						•		
R ₇	H	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•	•		•	•	•	•		•		•
	Cl										•						•			•						•		•	
R ₈	CH ₃	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•			•
	H						•					•			•												•	•	
R ₉	Isobutyl	•	•	•	•	•	•				•	•	•	•	•	•	•	•	•	•	•	•				•	•		•
	<i>n</i> -Propyl																						•	•			•		
	Isopropyl							•																					
	<i>sec</i> -Butyl								•																•				

^aBiosynthetic subunits are distinguished by color: unit A (red); unit B (blue); unit C (green); unit D (black). These data were assembled according to previous studies (1, 3, 4).

p-glycoprotein pumps and are active against multidrug-resistant tumor cell lines (6). These properties led to the advancement of cryptophycin 52 (LY355703), a synthetic analogue, to phase II clinical trials. In an initial study, dose-limiting toxicities of that analogue restricted its further advancement (8). However, in a subsequent phase II clinical trial performed on patients with

platinum-resistant ovarian cancer, the study concluded that the considerable rate of disease stabilization suggests that LY355703 might warrant further investigation (9). Moreover, another generation of **1** analogues has been synthesized that shield the reactive epoxide as a chlorohydrin or glycinate ester, resulting in improved solubility (10). Preclinical studies with the most promis-

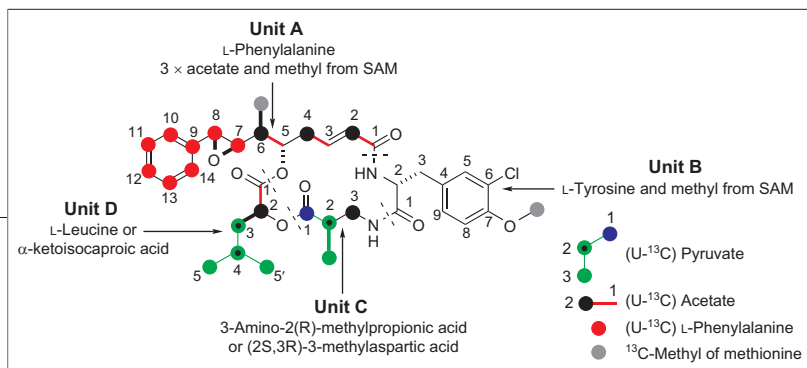


Figure 1. Summary of precursor incorporation experiments and labeling patterns to determine the origin of units A, B, C, and D of cryptophycin 1.

ing of these compounds show a dramatic increase in activity against a variety of tumors (10).

Because of the lack of large scale fermentation methods for isolation of cryptophycins, total synthesis was required to obtain adequate supplies for clinical evaluation (11). Several effective synthetic approaches have been developed for cryptophycin 3, a natural desoxy analogue of **1**. However, the most significant challenge has been the late-stage stereospecific installation of the epoxide moiety across the unit A C7–C8 double bond (Table 1), which is necessary due to the labile nature of this functionality. Initial efforts resulted in no better than a 3:1 mixture of **1** and the α -epoxide diastereomer that was difficult to separate (12). A more effective strategy to control epoxide stereochemistry was accomplished using Evans asymmetric aldol chemistry (13). This problem was addressed in an alternative manner for the synthesis of cryptophycin 52, which employed Sharpless methodology to generate a C7–C8 *syn* diol at the seco-cryptophycin stage and subsequent conversion to the β -epoxide (14, 15). More recently, we reported a convergent, chemoenzymatic synthesis of desoxy-deschloro-cryptophycin 1 by assembling a set of seco-cryptophycin chain elongation intermediates as the *N*-acetyl cysteamine (SNAC) ester and converting them to the corresponding cyclic depsipeptide using a cryptophycin (Crp) thioesterase (TE) mediated approach (16). Completion of the final step remained dependent on successful development of the predicted CYP450 derived from the native Crp biosynthetic system to install the β -epoxide functional group.

Here we provide a complete analysis of the biosynthetic origin and assembly for the A–D subunits comprising the Crp pathway through use of stable isotope precursor-labeling studies and the cloning, sequencing, and biochemical characterization of the Crp metabolic system. A comparative secondary metabolomic analysis was performed to localize and identify the 40 kb *crp* mixed polyketide synthase (PKS)/non-ribosomal peptide synthetase (NRPS) gene cluster. Annotation provided a detailed view of the collinear metabolic system

and revealed a number of unique enzymatic steps involved in subunit assembly, peptolide ring elaboration, and post-PKS/NRPS tailoring reactions. Precursor-directed biosynthesis using unnatural starter units afforded a suite of novel cryptophycins, providing direct evidence for the

flexibility of the biosynthetic enzymes along the assembly line. Finally, we demonstrate that CrpE is the cryptophycin CYP450 that provides an efficient *in vitro* method to generate the β -epoxide with complete regio- and stereochemical control. This enabled a novel chemoenzymatic synthesis of cryptophycin 2 using a tandem macrocyclization/epoxidation reaction sequence with CrpD TE and CrpE.

RESULTS AND DISCUSSION

Identification, Cloning, and Sequencing of the Cryptophycin Biosynthetic Gene Cluster: Comparative Cyanobiont Secondary Metabolome Analysis. The subunit structures (units A–D) comprising **1** (Figure 1) suggest an assembly from carboxylic acid and amino acid precursors by a mixed PKS/NRPS system. Degenerate polymerase chain reaction (PCR) primer sets are widely used to amplify and clone DNA fragments encoding segments of NRPS adenylation (A) (amino acid selecting) domains and ketosynthase (KS) (condensing enzyme) domains of type I PKSs (17, 18). Therefore, our initial strategy was to amplify A and KS domain DNA fragments and use the mixture of amplicons to detect cosmids containing PKS and NRPS genes from a *Nostoc* sp. ATCC 53789 genomic library. In the case of *Nostoc* sp. ATCC 53789, an unusually large number of nonoverlapping clones containing both NRPS and PKS genes were isolated reflecting a highly complex secondary metabolome. Four nonoverlapping PKS and NRPS gene-containing cosmids were partially sequenced to reveal distinct metabolic systems, but none were consistent with a Crp pathway (data not shown). Interestingly, three of the four cosmid DNA sequences revealed high similarity to secondary metabolic gene clusters found within the sequenced genome of the cycad symbiont, *Nostoc punctiforme* ATCC 29133 (data not shown) (19).

Nostopeptolides are the only previously described secondary metabolites produced by *N. punctiforme* ATCC 29133 (20), but there are no reports of its ability to produce cryptophycins. A Crp-sensitive bioassay using *Cryptococcus neoformans* (2) as an indicator strain confirmed the lack of Crp production by *N. punctiforme*

(data not shown). Therefore, we considered that a bioinformatics approach using A and KS domain sequences from *N. punctiforme* compared to A and KS domain sequences from *Nostoc* sp. ATCC 53789 would identify PKS and NRPS pathways unique to this lichen symbiont. Of the *Nostoc* sp. ATCC 53789 A domain sequences cloned and analyzed, only six were absent from *N. punctiforme*. One of the six DNA fragments contained within cosmid pNAM123 had a predicted aromatic amino acid specificity code (Asp235, Ala236, Ser239, Thr278, Ile299, Ala301, Gly322, Ile330) (21) and was selected as the candidate Crp pathway A domain for unit B. Significantly, PCR primers designed from the pNAM123 DNA insert (see Methods) generated an amplicon from *Nostoc* sp. GSV 224 genomic DNA whose sequence was 98% identical to the pNAM123 insert (data not shown). As further evidence that the Crp biosynthetic gene clusters from *Nostoc* sp. GSV 224 and *Nostoc* sp. ATCC 53789 are virtually identical, a fosmid was cloned from *Nostoc* sp. GSV 224 that partially contained *crpD* and encompassed all of the downstream genes associated with Crp production. The overlapping DNA sequences from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 were >99.9% identical and contained *crpE*, *crpF*, *crpG*, and *crpH* in the same order (Figure 2). Interestingly, the DNA sequences from both species were essentially identical up to the transposase region, after which they diverged completely, suggesting that the *crp* gene clusters were integrated at different genomic loci within the two species. These data also provide evidence that *crpH* represents the terminus of the Crp biosynthetic gene cluster. Probing a *Nostoc* sp. ATCC 53789 cosmid library provided pDHS500, whose insert was sequenced and found to contain NRPS and PKS genes consistent with the predicted architecture for the 3'-half of the Crp pathway. Additional library probing using PCR-based screening provided pDHS501 that overlapped with pDHS500 and contained the complete upstream PKS gene portion of the putative *crp* cluster.

The *crp* gene cluster (40,304 bp) is flanked by transposases and inverted repeats (interestingly, the curacin A and jamaicamide cyanobacterial biosynthetic gene clusters (18, 22) are also flanked by transposases) that appear to represent the limits of the metabolic system (Figure 2). The first open reading frame (ORF) within *crp* is a type I PKS gene (8823 bp) that is followed by a second modular PKS gene (10,407 bp) and then two modular NRPS genes (5829 and 10,029 bp) designated

crpA-D. A series of ORFs (*crpE-H*) downstream of the Crp PKS and NRPS genes are predicted to encode enzymes that catalyze functional group modifications (e.g., epoxidation (*crpE*) and chlorination (*crpH*)) present in **1** and many of its analogues. The *crp* biosynthetic gene cluster architecture (Figure 2) and a summary of each deduced protein sequence and corresponding functional role are summarized in Supplementary Table 1.

Proposed Assembly of the Cryptophycins. Biosynthesis of Type I PKS-Unit A: δ -Hydroxy-phenyloctenoic Acid.

As a first step toward unraveling the biosynthetic origin of unit A in the cryptophycins, sodium [1,2- $^{13}\text{C}_2$]acetate was provided in precursor incorporation studies to *Nostoc* sp. GSV 224 (Figure 1 and Supplementary Experiment A). In this experiment the precursor was diluted with unlabeled acetate to minimize the formation of interconnected [1,2- $^{13}\text{C}_2$]acetate units. The ^1H -decoupled ^{13}C NMR spectrum of labeled **1** (Figure 1) appeared as a 1:0.64:1.2 cluster of peaks (1.8%) for the signals of six contiguous carbons in unit A, specifically C1 to C6, which was consistent with the incorporation of three intact acetate units. Each triplet was composed of a singlet for the natural abundance ^{13}C flanked by peaks of a doublet for the incorporated ^{13}C . The level of ^{13}C enrichment (integration of doublet/integration of singlet) averaged 1.1%. The coupling constants associated with the doublets rigorously established that the three intact acetate units had been assimilated into C1–C2, C3–C4, and C5–C6. No ^{13}C enrichment was observed in any of the other carbon signals for unit A. Similar results were obtained when sodium [U- $^{13}\text{C}_3$]pyruvate was provided as a precursor to the cyanobacterium.

To determine the fate of the acetate protons in the formation of the C1–C2–C3–C4–C5–C6 segment of unit A, sodium [2- ^{13}C , $^2\text{H}_3$]acetate was provided as a precursor to the bacterial cells (Supplementary Experiment B). The ^2H -decoupled ^{13}C NMR spectrum of the labeled **1** exhibited enhanced signals (1.8%) for C2 and C6 and isotopically shifted ^{13}C signals for C4. The C4 signals were found in a 1:0.6:1.2 triplet at 36.70, 36.35, and 36.02 ppm assigned to undeuterated, monodeuterated, and dideuterated C4, respectively. About 80% of the deuterium on the ^{13}C incorporated into C4 had been retained, whereas all of the deuterium on the ^{13}C incorporated into C2 and C6 had been lost.

To establish the origin of oxygen atoms attached to C1 and C5 in unit A, sodium [1- ^{13}C , $^{18}\text{O}_2$]acetate was provided to a culture of *Nostoc* sp. GSV 224 (Supple-

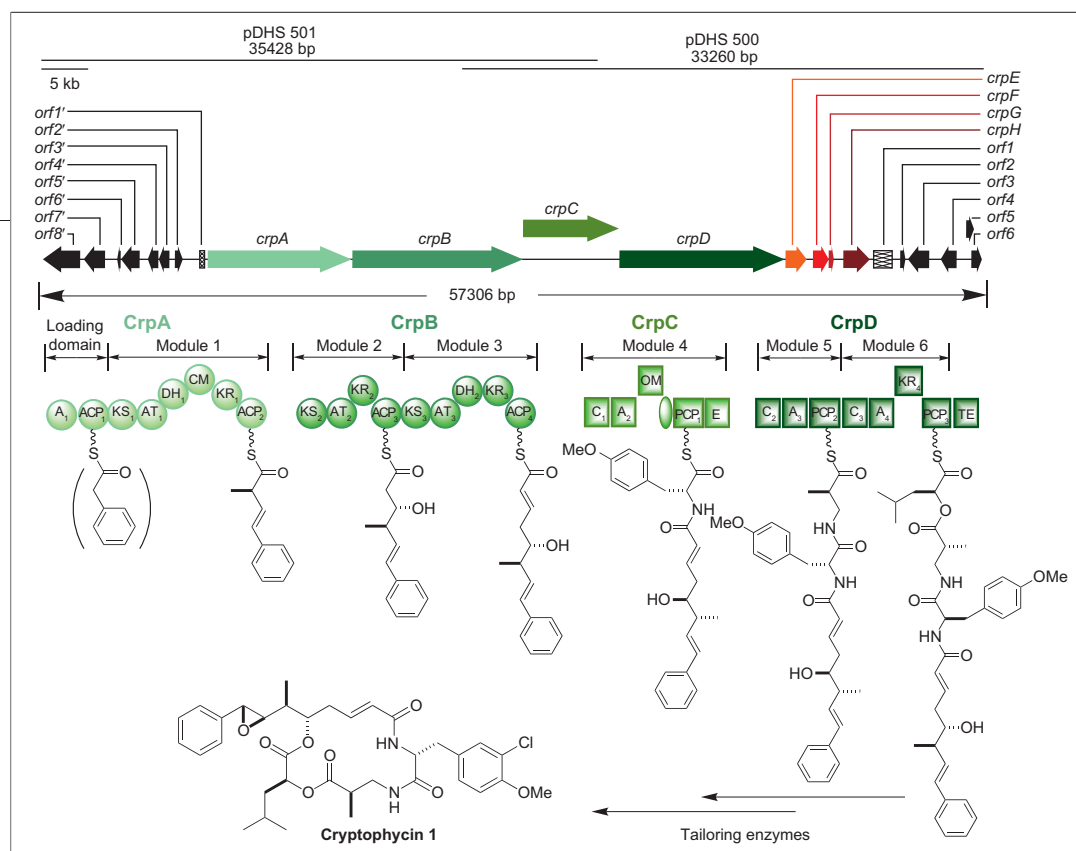


Figure 2. The Crp gene cluster and the deduced PKS/NRPS assembly line. Block arrows indicate ORFs. Genes that are cross-hatched are truncated ORFs or presumed pseudogenes. Genes depicted in green colors indicate that the corresponding gene products are predicted to function as NRPSs or PKSs. Genes shown in orange to red are predicted to encode tailoring enzymes or enzymes involved in precursor biosynthesis. Genes depicted in black are predicted to fall outside of the Crp biosynthetic gene cluster. Domains found within the proposed Crp PKS and NRPS assembly line are shown in circles (PKS) and squares (NRPS) with domain abbreviations (AT, acyltransferase domain; KS, ketosynthase domain; CM, C-methyltransferase domain; DH, dehydratase domain; KR, ketoreductase domain; ACP, acyl carrier protein; A, adenylation domain; C, condensation domain; PCP, peptidyl carrier protein; OM, proposed O-methyltransferase domain; E, epimerase; TE, thioesterase domain) denoting their function.

mentary Experiment C). The ^{13}C NMR spectrum of the labeled **1** showed ^{13}C peaks (1.3% enrichment) for C1 and C5 that were isotopically shifted upfield from the natural abundance lines by 0.03 and 0.04 ppm, respectively. This result revealed that essentially all of the ^{18}O on the ^{13}C incorporated into C1 and C5 from this precursor had been retained.

A precursor incorporation experiment with L-[methyl- ^{13}C]methionine established that the methyl group on C6 of unit A originates from the C_1 pool (Figure 1 and Supplementary Experiment D). The ^{13}C NMR spectrum showed a substantial enhancement (15%) of the carbon signal at δ 13.5.

The origin of the remaining carbons in unit A (i.e., C7, C8, and those in the phenyl group) were shown to be phenylalanine-derived from a precursor incorporation experiment with L-[U- $^{13}\text{C}_9$ - ^{15}N]phenylalanine (Supplementary Experiment E). The precursor was provided to *Nostoc* sp. GSV 224, and the bacterial culture was harvested and processed for **1**. The proton-decoupled ^{13}C NMR spectrum of **1** exhibited doublet and double doublet resonances (2.3% enrichment from the precursor)

for C7 and C8, respectively, due to ^{13}C enrichment from the precursor, with the natural abundance ^{13}C superimposed at the center of the carbon signal (other carbons appeared as multiplets). The direct ^2H NMR analysis of **1** obtained following supplementation of the cultures with L-[$^2\text{H}_8$]phenylalanine showed poor but positive incorporation of the precursor as indicated by weak deuterium signals for the phenyl group (Supplementary Experiment F). No significant level of deuterium on the benzylic epoxymethine

carbon, however, was detected. On the basis of these precursor incorporation results, phenylacetyl-CoA is a potential starter unit for the biosynthesis of unit A in **1** with chain elongation proceeding by a sequential addition of three malonate units (Figure 1 and Figure 3, panel a). However, feeding experiments with labeled phenylacetic acid and the corresponding N-acetylcysteine thioester failed to show incorporation. Although it might not be possible for these precursors to be transported across the cyanobacterial cell membrane, it is conceivable that another starter unit is involved (see below).

Recently, the origin of the starter unit of the cyanobacterial toxin microcystin (Mcy) was investigated by testing *in vitro* the activation and acylation of a series of candidate aryl and amino acids by McyG loading module (23). Acceptable substrates for the A loading domain were phenylpropanoids with McyG showing the highest selectivity toward *trans*-cinnamic acid. Interestingly, the CrpA A domain amino acid specificity code is identical to the corresponding McyG A domain, and thus, it is possible that *trans*-cinnamic acid functions as the Crp starter unit since it is also derived from phenyl-

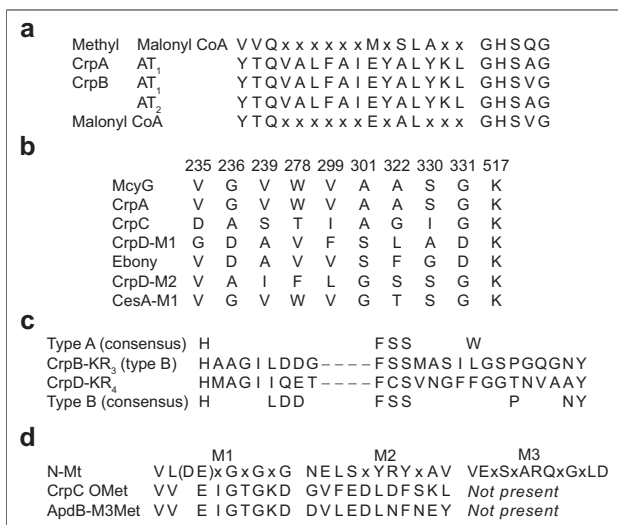


Figure 3. Sequence alignments between key conserved amino acid motifs of previously characterized PKS and NRPS domains, compared to deduced amino sequence of motifs identified from Crp PKS and NRPS domains.

a) Identification of AT specificity based on consensus sequences motifs from malonyl-CoA and methylmalonyl-CoA AT domains. **b)** Alignment of the proposed active site region of KR domains CrpA-KR₃ and CrpD-KR used previously to classify KR domains as type A and type B (24). Residue numbering is relative to residue 1 of the EryA KR domain (54). **c)** Comparison of conserved motifs from NRPS *N*-methyltransferase domains (N-Mt domains) (27) with the CrpC and anabaenopeptidase (ApdB)-M3 methyltransferase domain motifs. **d)** The deduced codes of the Crp adenylation domains (CrpA-phenylacetate-like starter unit; CrpC, chloro-*L*-tyrosine, CrpD-M1, methyl β -alanine; CrpD-M2, α -ketoisocaproic acid) as well as NRPS codes for the Ebony protein (*Drosophila melanogaster*), β -alanine; McyG (microcystin PKS), proposed phenylacetate specific A domain; CesA-M1 (cereulide NRPS), α -ketoisocaproic acid specific A domain.

alanine (Figure 3, panel b). A remaining mystery is the mechanistic basis for an unprecedented one-carbon truncation within the growing unit A PK chain derived from a phenylpropanoid starter unit.

The precursor incorporation studies described above are consistent with the established three-module domain organization of CrpA (A-ACP-KS-AT-DH-CM-KR-ACP) and CrpB (KS-AT-KR-ACP-KS-AT-DH-KR-ACP) that are collinear with the **1** unit A structure (Figure 2). Several structural variations that deviate from the collinearity rule in the CrpA/B PKS are noteworthy. One is the smaller macrocycle (14-membered ring) found in cryptophycin 26 (Table 1), which is unique compared to all other cryptophycins. Formation of cryptophycin 26 requires that the C3 hydroxyl group serves as the nucleophile for CrpD TE mediated ring closure (16). Bypassing of the CrpB dehydratase (DH) might preserve the C3 hydroxyl group in cryptophycin 26 as well as cryptophycin 30 (Table 1). In another natural analogue (cryptophycin 327) a C-2/C-3 *cis*-double bond occurs instead of the C-2/C-3 *trans*-double bond common to all other cryptophycins. The CrpB KR₂ domain is predicted to be responsible for C3 ketoreduction, and its amino acid sequence is coincident with an *S*-specific KR (or type B) based on bioinformatics analysis (24–26) (Figure 3, panel c), consistent with the C-2/C-3 double bond. Presumably the CrpB KR₂ domain also generates an *R* C-3 stereoisomer at a low level, with the CrpB DH₂ catalyzing dehydration to the C-2/C-3 *cis* olefin.

NRPS: Incorporation of Unit B and Unit C in the Cryptophycins. Unit B: 3-Chloro-*O*-methyl-*D*-tyrosine.

Unit C: Methyl- β -alanine. Unit B of **1** is 3-chloro-*O*-methyl-*D*-tyrosine, but other rare cryptophycins incorporate *L*-amino acids (e.g., *L*-tyrosine), dichlorotyrosines (e.g., 3,5-dichloro-*O*-methyl-*D*-tyrosine), or desmethyl-deschlorotyrosines (e.g., *O*-methyl-*D*-tyrosine, *D*-tyrosine) (Table 1). The origin of unit B in **1** was explored using *L*-[1-¹³C]tyrosine and *D*- and *L*-3-([4-¹³C]-phenyl)-alanine precursor incorporation studies (Supplementary Experiments G and H, respectively). The ¹³C NMR spectrum from **1** produced in the presence of *L*-[1-¹³C]tyrosine showed a 4% enhancement of C-1 of unit B. Feeding experiments with *D*- and *L*-3-([4-¹³C]-phenyl)-alanine resulted in deuterium-enriched **1**, indicating that unit B is derived from *L*-tyrosine, and the Crp NRPS catalyzes the incorporation of *O*-methyl-*D*- and *L*-tyrosine (i.e., *L*-3-([4-¹³C]-phenyl)-alanine) (Figure 1).

CrpC is a monomodular NRPS containing an elongation module with an A domain that has an NRPS amino acid specificity code (Figure 3, panel a) most similar (75% identical, 87% similar) to the ApdB-M4 NRPS A domain (*O*-methyl-*L*-tyrosine, or *N,O*-dimethyl chloro-*L*-tyrosine). The CrpC A domain, bears an *S*-adenosylmethionine (SAM)-dependent methyltransferase between motifs A8 and A9 (27) of the polypeptide (Figure 2). Protein database searching revealed that the most similar peptide sequence (64% identity, 84% similarity) to the CrpC methyltransferase domain (~400 amino acids) is the ApdB-M3 methyltransferase of the

Apd NRPS (28). We propose the CrpC and ApdB-M3 methyltransferases (*O*-methylation of tyrosine) are a new type of NRPS domain that specify *O*-methylation, which is supported by unique amino acid sequence motifs (Figure 3, panel d). Both the ApdB-M3 and CrpC methyltransferase domains lack a core *N*-methyltransferase amino acid sequence motif (M3), and residues within core methyltransferase motifs M1 and M2 are different. Following the presumed OMet domain are A domain motifs (A9 and A10) of unknown function, a PCP domain, and an epimerase (E) domain (27) that is predicted to catalyze the conversion of L-tyrosine to D-tyrosine. Unit B is structurally modified by mono- or dichlorination of the tyrosine subunit (Table 1). It remains to be determined whether halogenation occurs at the subunit stage to generate a precursor pool of chlorotyrosine or during a later stage of natural product assembly.

CrpH, whose gene (*crpH*) represents the presumed terminus of the *crp* gene cluster is the predicted halogenase responsible for modification of tyrosine at the position ortho to the OMet moiety (Supplementary Table 1). Database comparisons with the deduced CrpH peptide sequence revealed that it resembles a number of previously described non-heme flavin-dependent halogenases, including those from the rebeccamycin (29) and pyoluteorin (30) pathways. We expect that CrpH is responsible for mono- and dichlorination of the aromatic ring of unit B for the various Crp natural products.

Unit C of **1** is methyl-β-alanine, whereas in a small number of analogues (*e.g.*, cryptophycins 18, 24, 29, 176, 326; Table 1) β-alanine is incorporated for the elongation step. In primary metabolism, phosphopantetheine formation requires β-alanine that is in turn derived from aspartate *via* a PanD decarboxylase (31). To determine if methyl β-alanine is formed from methyl-aspartate (MASP), a MASP synthesis (32) was developed using [¹³C]iodomethane instead of iodomethane to provide a 1:1 (*SR/SS*) diastereomeric mixture of the compounds (Supplementary Experiment I). This enabled precursor incorporation studies using a ¹³C-labeled methyl group in cultures of *Nostoc* sp. GSV 224. A 1.4% enrichment of ¹³C label from [methyl-¹³C]-(2*S*,3*R/S*)-3-MASP in the methyl group of unit C in **1** demonstrated that unit C is derived from this biosynthetic subunit (Figure 1). In addition, feeding studies using [U-¹³C]pyruvate resulted in intact incorporation of pyruvate into the methyl-β-alanine unit (Figure 1 and Supplementary Experiment J).

The only other natural product that is known to contain methyl-β-alanine is vicienistatin. However, a combination of feeding studies and analysis of the DNA sequence of the entire vicienistatin biosynthetic cluster (33) indicates a unique origin for this subunit. Specifically, the methyl-β-alanine unit is presumably derived from glutamate that is rearranged by a glutamate mutase to form MASP and finally decarboxylated by a pyridoxal-5'-phosphate-dependent decarboxylase (33, 34). Intact incorporation of pyruvate by the Crp biosynthetic pathway is inconsistent with the formation of MASP derived by rearrangement from glutamic acid and suggests that MASP in *Nostoc* may originate from a biosynthetic pathway analogous to branched chain amino acid biosynthesis. The primary sequence of CrpG is similar to various PanD enzymes (Supplementary Table 1). Interestingly, when ¹³C-labeled *S*-methyl-β-alanine was fed to cultures of *Nostoc* sp. GSV 224, it was incorporated with a percent yield comparable to the incorporation of the *R*-isomer (Supplementary precursor-directed biosynthesis). It appears that the CrpD A2 domain does not discriminate between these diastereomeric forms of the subunit. This result suggests that CrpG displays high selectivity toward *R*-MASP to afford the 2*R*-stereoisomer of methyl-β-alanine. The ability of the CrpD-M1 A domain to activate methyl-β-alanine/β-alanine is predicted based on its sequence similarity to previously characterized (21) β-alanine A domains (Figure 3, panel b).

Mosaic NRPS/PKS Module. Unit D: Variation and Incorporation of Branch Chain α-Hydroxy Acids. On the basis of reasonable biosynthetic principles, it was evident that unit D of **1** is derived from α-hydroxyisocaproic acid. Interestingly, α-hydroxyisovaleric acid, α-hydroxyisobutyric acid, or α-hydroxyvaleric acid is also incorporated based on the structures of several natural Crp analogues (Table 1; cryptophycins 19, 21, 49, 50, 54, 326). We sought to determine if L-leucine is the direct precursor for unit D by conducting precursor incorporation experiments with D,L-[5-²H₃]leucine (Supplementary Experiment K) and with L-[1-¹³C]leucine (Figure 1, Supplementary Experiment L). In feeding experiments with D,L-[5-²H₃]leucine, **1** was isolated and shown to have a ²H resonance at 0.827 ppm (1.8% incorporation). Intact and equal incorporation (0.5%) of the methyl side chains derived from leucine into the unit D methyl groups was evident in the ¹³C NMR spectrum of **1** isolated from these *Nostoc* cultures. In feed-

ing studies with L-[1- ^{13}C]leucine, the ^{13}C NMR spectrum showed incorporation of 1% of the label from C-1 of the precursor into C-1 of unit D. Precursor incorporation studies with α -hydroxyisocaproic acid, however, did not result in labeled cryptophycins (Supplementary Experiment M). Unlike α -hydroxyisocaproic acid, α -ketoisocaproate is a common leucine biosynthesis intermediate (35), formed from leucine *via* transamination. The direct contribution of α -ketoisocaproate to unit D was tested by precursor incorporation studies with α -[1- ^{13}C]ketoisocaproic acid. **1** extracted from the resulting cultures was subjected to ^{13}C NMR analysis that showed enhancement (0.9%) of the ester carbonyl (Figure 1; Supplementary Experiment N). On the basis of the required conversion of leucine to the corresponding α -keto acid, we surmise that CrpF, a putative non-heme oxygenase (Supplementary Table 1), might catalyze oxidative deamination of leucine yielding α -ketoisocaproate. Final KR-mediated reduction (see below) would be required to generate the nascent unit D precursor.

The final module within the Crp PKS/NRPS is the second module of CrpD (CrpD-M6) that bears an NRPS-like elongation module ($\text{C}_3\text{-A}_4\text{-KR}_4\text{-PCP}_3$) and the terminal TE domain, CrpD TE (16) (Figure 2). Following presumed activation and transfer of α -ketoisocaproic acid to the CrpD-M6 PCP_3 , the tethered α -keto acid might serve as a substrate for the upstream KR-like domain. In polyketide biosynthesis, KR domains have been demonstrated to reduce β -keto groups with no reported precedent for α -keto reducing KRs. The valinomycin (*vlm*) (36) and cereulide (*ces*) (37) NRPSs contain dehydrogenases similar to that found in CrpD. One of the *ces* dehydrogenases was recently demonstrated to catalyze α -keto reduction of a thioester bound 2-keto isocaproyl-PCP substrate (38).

NRPS C domains typically function as amide synthetases, but the domain placement and predicted substrate of the CrpD C_3 domain suggest that it is an ester synthetase. Highly homologous domains from the *vlm* (36) and *ces* (37) NRPSs were also predicted to bear ester synthetase C domains. Recently the fumonisins-A PCP-C didomain was shown to catalyze this type of ester bond formation (39). Finally, the hydroxyl group of unit A and the leucic acid carboxyl group of unit D are linked to form a lactone, catalyzed by CrpD TE (16).

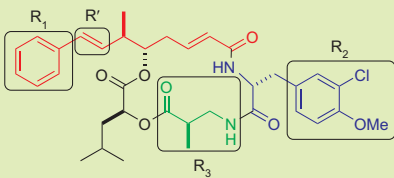
Precursor-Directed Biosynthesis of Novel Cryptophycins.

The unusually large number of natural product analogues generated by the Crp biosynthetic system suggests that it is comprised of flexible enzymes for depsipeptide assembly and tailoring. We sought to assess further the versatility of the Crp assembly line using precursor-directed biosynthesis. A series of amino acids and halogens were used in order to probe the flexibility of enzymes responsible for incorporation of units A, B, and C.

We first targeted the aromatic groups of unit A as this is among the least varied in the natural cryptophycins (Table 2), yet synthetic chemical structural modification to the aromatic group can provide enhanced bioactivity (11). Our demonstration that phenylalanine is converted into a substrate for the CrpA A_1 domain suggested that analogues of this amino acid might provide insight into CrpA initiation domain flexibility. As an initial experiment, we used *p*-methylphenylalanine for precursor incorporation, which led to the new metabolite cryptophycin 111 (Table 2 and Supplementary Information). Further experiments included a diverse set of phenylalanine analogues, including those with larger *p*-alkyl chains (*i.e.*, ethyl and alkyl hydroxyl) and more bulky ring systems (naphthyl and biphenyl), cyano-appended rings, and para-substituted halides (Cl, Br, F, and I) (Table 2). In each case, new cryptophycins were generated with expected starter units, revealing that the Crp loading domain is remarkably flexible. Interestingly, many of these amino acids were also incorporated within unit B (Table 2). Moreover, this work demonstrates that the Crp PKS/NRPS is able to channel and process alternative substrates of varying structural complexity (Table 2). In most cases, the new compounds obtained were modified by the presumed CrpE CYP450, revealing significant tolerance of the epoxidase toward unnatural substrates (Table 2).

The *in vivo* specificity of the CrpH halogenase was also investigated by adding CaBr_2 and CaI_2 to the cyanobacterial culture medium. Surprisingly, both bromine and iodine were incorporated into unit B, indicating an additional opportunity for *in vivo* generation of diverse products (Table 2). This was particularly intriguing in view of recent reports that replacement of chlorine for iodine has not been observed previously in natural product halogenases (40). Interestingly, the epoxidized product of the brominated molecule (Table 2; cryptophycin 104) was extracted from the bacterial culture, but the

TABLE 2. Summary of the precursor-directed biosynthesis to assess substrate tolerance in the Crp pathway^a



Precursors	Cryptophycin products				
	R'		R ₁	R ₂	R ₃
	<i>trans</i> -Styrene	β -Epoxide			
D,L-4-Fluorophenylalanine	110	115	4-Fluorophenyl	*	*
D,L-4-Chlorophenylalanine	124	125	4-Chlorophenyl	*	*
	324		4-Chlorophenyl	4-Methoxyphenyl	*
D,L-4-Bromophenylalanine	304	305	4-Bromophenyl	*	*
D,L-4-Iodophenylalanine	310		4-Iodophenyl	*	*
L-4-Methylphenylalanine	111	117	4-Methylphenyl	*	*
	312		4-Methylphenyl	4-Methoxyphenyl	*
L-4-Ethylphenylalanine	339		4-Ethylphenyl	*	*
L-4-Trifluoromethylphenylalanine	161		4-Trifluoromethylphenyl	*	*
L-4-Cyanophenylalanine	318		4-Cyanophenyl	*	*
D,L-3-Fluorophenylalanine	311	303	3-Fluorophenyl	*	*
(2 <i>R</i>)-3-Fluorotyrosine	211	316	*	3-Fluoro-5-chloro-4-methoxy-phenyl	*
(2 <i>R</i>)-3-Fluoro-4-[² H ₃]-methoxytyrosine	210	190	*	3-Fluoro-4-[² H ₃]-methoxyphenyl	*
L-3-Iodotyrosine	320	319	*	3-Iodo-4-methoxyphenyl	*
D,L-2-Fluorophenylalanine	181	182	2-Fluorophenyl	*	*
(2 <i>R</i>)-3-(3,4-[² H ₂]-Methylenedioxy-phenyl)alanine	208	209	*	3,4-[² H ₂]-Methylenedioxy-phenyl	*
(2 <i>R</i>)-3-(3-Methyl-4-[² H ₃]-methoxy-phenyl)alanine	189		*	3-Methyl-4-[² H ₃]-methoxy-phenyl	*
L-4-Methylhydroxyphenylalanine	336		*	4-Methylhydroxyphenyl	*
	238		4-Methylhydroxyphenyl	3-Chloro-4-methoxyphenyl	*
L-2-Naphthylalanine	334		*	Naphthyl	*
	172		Naphthyl	*	*
	335		Naphthyl	Naphthyl	*
CaBr ₂	315	104	*	3-Bromo-4-methoxyphenyl	*
		313	*	3-Bromo-4-hydroxyphenyl	β -Alanine
		314	*	3-Bromo-4-hydroxyphenyl	*
CaI ₂	320	319	*	3-Iodo-4-methoxyphenyl	*
β -Alanine	29	21	*	*	β -Alanine
	24		*	4-Methoxyphenyl	β -Alanine
(2 <i>S</i>)-3-Aminoisobutyric acid	342	343	*	*	(2 <i>S</i>)-3-Aminoisobutyric acid
<i>gem</i> -Dimethyl- β -alanine	52		*	*	<i>gem</i> -Dimethyl- β -alanine

^aThe asterisks indicate that the R group is the same as the R group of the illustration.

iodinated form (Table 2; cryptophycin 320) was also isolated as the *trans*-styrene, indicating that unit B might impose structural constraints on CrpE substrate selectivity.

Flexibility of the Crp pathway unit C enzymes was investigated by evaluating precursor incorporation using β -alanine, and (2*S*)-3-aminoisobutyric acid. Both of these amino acids were incorporated into unit C with complete downstream processing to the expected cyclic depsipeptide structures (Table 2; cryptophycins 24, 29,

342, 343). Cryptophycin 52, the synthetic lead molecule that was advanced to clinical trials, is identical to **1** except for a *gem*-dimethyl β -alanine as unit C. On the basis of precursor incorporation studies showing that the CrpD A domain is tolerant of unnatural subunits, we sought to test whether the Crp biosynthetic machinery was capable of forming cryptophycin 52 *in vivo*. Direct supplementation of *gem*-dimethyl β -alanine to *Nostoc* cultures resulted in a product that was shown by NMR to

be cryptophycin 52 (Table 2 and Supplementary Information for NMR characterization of cryptophycins). This is the first report using precursor-directed biosynthesis to produce cryptophycin 52, further advancing the potential of the Crp biosynthetic machinery to access important bioactive Crp analogues. Previous to this work, access to this important analogue was only possible by total synthesis (41).

Chemoenzymatic Synthesis of Cryptophycin 2:

Sequential Synthesis from Seco-cryptophycin 4 Using CrpD TE/CrpE P450. Recently, we reported the ability of excised CrpD TE to catalyze *in vitro* macrocyclization of several acyl-peptide SNAC variants of Crp chain-elongation intermediates (16). One of the depsipeptide products obtained from the TE was desoxycryptophycin 2, the presumed substrate for the Crp epoxidase. The desoxycryptophycins have significantly reduced bioactivity, and it is installation of the β -epoxide that represents a significant challenge in Crp total synthesis. The Crp epoxidase was readily identified as the putative product of *crpE* whose deduced amino acid sequence is similar to numerous CYP450s involved in hydroxylation or epoxidation. To explore its specific function, the gene was cloned and overexpressed in *Escherichia coli* to provide a soluble histidine tagged/maltose-binding protein variant (data not shown). Incubation of CrpE with cryptophycin 4 resulted in >75% conversion to a compound with a mass consistent with cryptophycin 2 (Figure 4). Scale-up provided sufficient quantities of the epoxidized material to confirm by ^1H NMR that the compound is a single diastereomer (*i.e.*, the β -epoxide) and identical in all respects to cryptophycin 2 (42). This result confirms that CrpE is the Crp epoxidase that enables efficient, stereospecific installation of a β -epoxide into the cryptophycins. As an additional demonstration of the flexibility of this system, we employed the seco-cryptophycin 4 intermediate in a tandem *in situ* reaction that included CrpD TE, CrpE, and all required cofactors (see Methods). This experimental system provided direct isolation of cryptophycin 2 as a single product without the need to purify the desoxycryptophycin 2 intermediate (Figure 4). Effective *in vitro* reconstitution of these two final biosynthetic steps promises immediate access to new cryptophycins in the search for improved anticancer lead compounds.

In summary, the cyanobacterial symbiont-derived *Nostoc* sp. ATCC 53789 and the highly related *Nostoc* sp. GSV 224 produce the cryptophycin natural products.

These compounds have shown significant potential as anticancer therapeutic agents, and a number of analogues are currently being explored as clinical candidates. In this report, we focused on five allied studies, including (i) cloning, sequencing, and bioinformatics analysis of the Crp gene cluster, (ii) precursor incorporation studies to reveal the specific source of each biosynthetic subunit, (iii) precursor directed biosynthesis to probe the flexibility of Crp metabolic enzymes involved in chain assembly, cyclization, and tailoring, (iv) identification and functional analysis of the CrpE CYP450 that catalyzes β -epoxide formation to yield **1** and related analogues, and (v) tandem *in vitro* assembly of cryptophycin 2 from seco-cryptophycin 4 using CrpD TE and CrpE reconstituted in a single reaction vessel.

Because of the remarkable number of secondary metabolic gene clusters associated with *Nostoc* sp. genomes, a comparative analysis was employed to survey the complete secondary metabolome of *N. punctiforme* and use it as a tool to identify the *crp* gene cluster in both *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224. To our knowledge, this is the first report of natural product gene cluster identification using a comparative metabolomic strategy, made possible by the availability of a full genome sequence in a highly related species. Previous studies have shown the close phylogenetic relationship between lichen cyanobionts (43), but our reported strategy probes further the potential value of investigating cyanobiont phylogeny and secondary metabolism.

A particularly significant characteristic of the Crp biosynthetic system is its ability to generate >25 related natural products that include alterations in subunit structures, including units A, B, C, and D, depsipeptide ring size (14 vs 16), and variations in tailoring reactions, including epoxidation and halogenation. Subunit A structural variants are presumably derived from imprecise PKS processing reactions, including Crp module 3 KR₃ and DH₂ domains. These events result in variant double bond configuration or presence of a second hydroxyl group at C-3 in unit A (Table 1). NRPS enzymes involved in assembly of units B, C, and D show flexibility in amino acid and α -hydroxy acid subunit selectivity, resulting in the incorporation of diverse, yet related, extender groups. Tailoring reactions involving the *crpE*-encoded halogenase can result in one or two chlorine atoms in the unit B aromatic ring of tyrosine. Precursor incorporation studies revealed definitively the source of

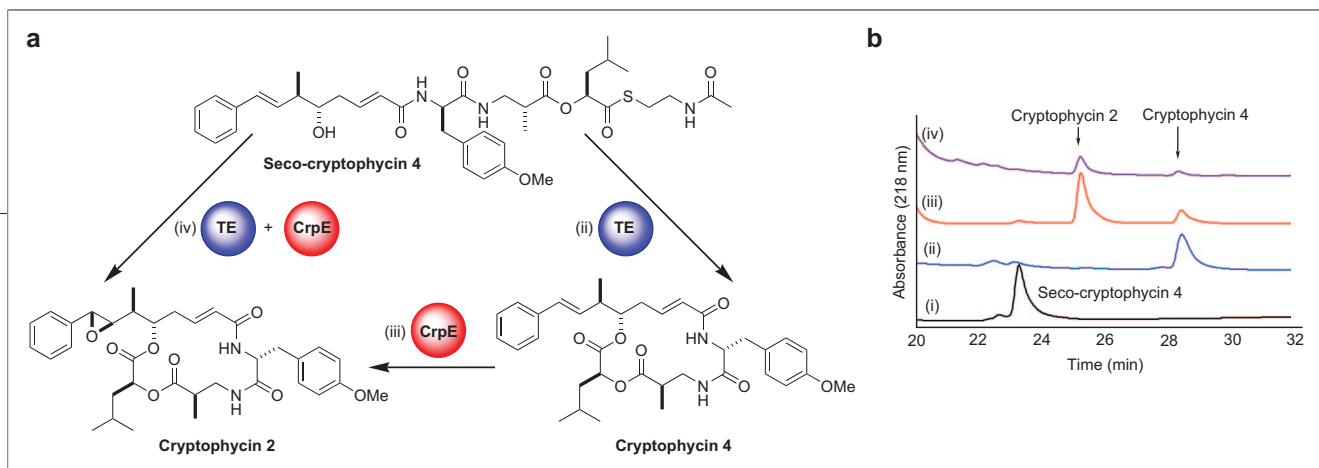


Figure 4. CrpD TE mediated cyclization and CrpE mediated epoxidation. **a)** Illustration of the reactions catalyzed by CrpD TE and CrpE. **b)** HPLC traces of: i, 10.0 μM SNAC-thioester of seco-cryptophycin 4 (black line); ii, CrpD TE catalyzed cyclization of 9.0 μM seco-cryptophycin 4 (blue line); iii, CrpE mediated epoxidation of 9.5 μM cryptophycin 4 (previously cyclized by CrpD TE and HPLC purified) to produce cryptophycin 2 (red line); iv, tandem (one reaction vessel) CrpD TE and CrpE mediated cyclization and epoxidation of 5.2 μM seco-cryptophycin 4 (purple line). The reactions contained 0.1 M Tris pH 7.4, 5% DMSO, 1.8 μM Crp TE and/or 3 μM CrpE with 100 $\mu\text{g mL}^{-1}$ ferredoxin, 0.2 unit mL^{-1} ferredoxin-NADP⁺ reductase, 1.4 mM NADPH, 10 mM glucose-6-phosphate, 8 units mL^{-1} glucose-6-phosphate dehydrogenase at 30 °C from 2 to 4 h.

units B, C, and D. Although unit A was shown to be derived from three acetate groups, the ultimate source of the “phenylacetate” starter group requires further investigation. Phenylalanine is the likely Crp pathway initiation group, but the precise steps involved in conversion to the “phenylacetate” equivalent remains unknown. Precursor-directed incorporation studies revealed additional flexibility for the Crp biosynthetic pathway through its ability to accept a number of struc-

turally diverse starter units that were channeled, processed, cyclized, and released as fully elaborated novel depsipeptide natural products. Finally, we demonstrated for the first time a stereospecific epoxidation reaction resulting in effective conversion of cryptophycin 4 to cryptophycin 2 using the CrpE CYP450. This unique biocatalyst, along with the versatile CrpD TE domain, provides new, useful tools for the chemoenzymatic synthesis of Crp anticancer agents.

METHODS

General Microbiology and Molecular Biology Procedures: Chemicals, Bacterial Strains, Culture Conditions, and DNA Subcloning.

Molecular biology reagents and enzymes were supplied by New England Biolabs except for Pfu (Stratagene), dNTPs (Takara), T4 DNA Ligase (Invitrogen), pSMART-HCKan (Lucigen), and Wizard SV PCR Clean-up kit (Promega). When necessary, other chemicals were purchased from Sigma-Aldrich. *Nostoc* cultures were routinely maintained on BG-11 agar plates prepared as previously described (44). *Nostoc* strains prepared for genomic DNA extraction and production of Crp were grown in BG-11 medium (44) at 28 °C with constant illumination (9 W m^{-2}) and aeration. The *E. coli* strains used for cloning and plasmid harvesting were XL-1 Blue (Stratagene) and TOP-10 (Invitrogen). Protein overexpression was performed in *E. coli* BL21 (DE3) (Invitrogen). All *E. coli* strains were grown in Luria-Bertani (LB) broth and when necessary supplemented with 25 $\mu\text{g mL}^{-1}$ kanamycin. Preparation and manipulation of plasmid DNA from *E. coli* were accomplished using standard methods (45). DNA sequencing was performed at the University of Michigan DNA Sequencing Core or the University of Minnesota Advanced Genetic Analysis Center using ABI Prism sequencers and the dye termination method.

Subtractive Analysis of *Nostoc* A and KS Domains. Amplicons encoding both KS domain (~750 bp) and A domain fragments (~1100 bp) were obtained using conditions and PCR primers described previously (46, 47). The resulting amplicons were cloned into the sequencing vector pCR2.1 (Invitrogen) and 96 clones bearing inserts for both KS and A domains were selected for DNA sequencing. Sequences of KS and A domains were compared by multiple sequence alignment using the Clustal X program (48) (version 1.82) with A and KS domains from the deduced sequence of NRPSs and PKSs encoded by ORFs identi-

fied within the *N. punctiforme* genome sequence (GenBank accession # NZ_AAA000000000).

Cloning, Sequencing, and Annotating the Crp Cluster. High molecular weight genomic DNA was harvested from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 according to established protocols (49) and also using a FASTDNA spin kit (Qbiogene). For cosmid library construction total *Nostoc* DNA was partially digested using *Sau3A*I, dephosphorylated, and size selected (36–40 kb) using a CHEF gel with a CHEF-DR III PFGE system (Biorad), followed by ligation into *Bam*HI-digested SuperCos1 (Stratagene). The ligation mixture was packaged using the Giga-pack III XL Packaging Extract Kit (Stratagene), and the resulting library was titrated and amplified according to manufacturer instructions. Fosmid libraries were constructed in an analogous fashion but without prior restriction enzyme digestion, and blunt-ended fragments were cloned directly into pCC1Fos fosmid according to manufacturer instructions (Epicentre). The insert within pNAM123 (encoding a portion of an A domain) was liberated by *Eco*RI digestion, and the resulting DNA fragment was radiolabeled using the RadPrime labeling kit (Pharmacia) with [α -³²P] dCTP (Amersham) according to manufacturer directions. The radiolabeled fragments were used to probe the genomic library using standard colony hybridization protocols (45). The cosmid pDHS500, identified from probing with the pNAM123 insert, was fully sequenced using a shot-gun cloning approach where pDHS500 was fragmented in a nebulizer (IPT Medical Products, Inc.) with the following parameters, 4.4×10^4 Pa of N_2 for 4.0 min. Likewise, the fosmid DNA from pDHS501 was nebulized and the 2–6 kb fragments collected. The resulting pDHS500 and 501 fragments were blunt-ended by Klenow and T7 DNA polymerase and phosphorylated by treatment with T4 kinase and ATP immediately prior to fractionation on a low melting agarose gel. Fragments in the 2–6 kb size range were eluted, concentrated, and used separately for ligation into

the *Sma*I site of pUC18 and pSMART-HCKan (Lucigen). The sequences of pUC18 and pSMART-HCKan inserts were trimmed to remove sequences of SuperCos-1 and pCC1Fos and assembled into contigs using the SeqMan 5.06 program of the DNASTAR software package. The pDHS501 fosmid was identified by a PCR-screening strategy employing primers designed from the *crpD* TE region (16) and a second set (5'-GCATTGTC-ATTCTGGTGAGGC-3' and 5'-CCTGCTGCTAAGGCTATTCCAAG-3') for a portion of *crpB* contained within pDHS500. Two rounds of PCR were used to identify clones that produced amplicons using the PKS primers, but not the TE domain primers. The first round of PCR was used to identify plates that contained amplicons of the desired size. pDHS500 was used as a positive control for both primer sets. Ten-microliter portions of each culture from a given 96-well plate were pooled. The DNA was then used in PCR reactions using the TE and the PKS primer sets. The PCR reactions contained 1 μ L of pooled DNA, 1 μ M of each primer, 200 μ M dNTP, 5 μ L of Taq polymerase buffer, and 0.5 μ L of Taq DNA polymerase, and water was added to a final volume of 50 μ L. The PCR reactions were run at 94 °C for 2 min and cycled 30 times with the program: 94 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min. A *crpB* PKS amplicon was generated from one of the clones with the *crpB* primer set but not with the *crpD* TE primer set. The clone (pDHS501) producing this amplicon was identified and shown to extend from the 5' end of pDHS500. ORFs from pDHS500 and pDHS501 were identified using Vector NTI software package version 10 and their deduced amino acid sequences compared to others in the GenBank database. The domain arrangements of the Crp PKS and NRPS proteins were determined using the web-based software program PKS-NRPS (50). The PKS-NRPS program was also used to determine the predicted specificity codes of the Crp NRPS A domains. The *crp* gene cluster DNA sequence has been deposited in GenBank under accession number EF159954.

Overexpression and Purification of CrpE. The CrpE gene was amplified by PCR using cosmid pDHS500 as template. A typical 50 μ L reaction mixture contained 5 ng of pDHS500, 2 μ M forward primer (5'-TGC GGA TCC ATG ATT AAT ACT GCT AAA TCC-3') and 2 μ M reverse primer (5'-ACG CGA ATT CTT ACA ATA CAA CCA TTT TTA ATC C-3') (*Bam*HI and *Eco*RI sites are underlined), 200 μ M dNTP, 5 μ L of 10 \times PCR buffer. Conditions for *crpE* amplifications included an initial 5 min denaturation step (94 °C) and cycling conditions of 94 °C (30 s), 58.5 °C (30 s), and 72 °C (1 min 45 s) for 35 cycles followed by a final 72 °C extension step (7 min). The *crpE* amplicon was phosphorylated using T4 kinase and ligated into pSMART-HCKan to produce pDing1. The insert within pDing1 was sequenced and shown to be free of PCR errors. pDing1 was digested with *Bam*HI and *Eco*RI and ligated into linear pSJ8 (*Eco*RI/*Bam*HI cut). The resulting plasmid, pDing2, and chaperone expressing plasmid pGRO7 were used to cotransform *E. coli* BL21(DE3) to ampicillin (amp) (50 μ g mL⁻¹) and chloramphenicol (chl) (25 μ g mL⁻¹) resistance. A 5 mL overnight culture was diluted in 1 L of LB supplemented with amp (50 μ g mL⁻¹), chl (25 μ g mL⁻¹), 0.25 mM Fe(NH₄)₂(SO₄)₂, 1 mM thiamine, and 0.25 mM of 5-aminolevulinic acid that was added 30 min prior to induction (OD₆₀₀ ~ 0.6). The culture was then cooled to 4 °C before inducing with 0.1 mM Isopropyl- β -D-thiogalactopyranoside and 1 g L mL⁻¹ of L-arabinose. The culture was grown at 15 °C with constant shaking (200 rpm) for 20 h. The cells were pelleted by centrifugation and resuspended in 80 mL of PBS buffer (140 mM NaCl, 2.7 mM potassium chloride, 10 mM sodium hydrogen phosphate, and 1.8 mM potassium dihydrogen phosphate, pH 7.4, 3 mM β -mercaptoethanol, 10% glycerol). The cell suspension was then sonicated and the lysate collected following centrifugation (40,000g for 45 min). The resulting supernatant was collected and incubated with pre-equilibrated amylase agarose

resin at 4 °C for 3 h with agitation. The amylase agarose resin was washed (100 column volumes of lysis buffer) and MBP-His-CrpE eluted with lysis buffer containing 8 mM maltose. Maltose was removed from the protein sample with a PD-10 column, and the MBP-His-CrpE fusion protein was then treated with His-TEV protease at 4 °C overnight to remove the MBP-His portion. The CrpE protein was separated from the MBP-His polypeptide by passing the mixture through a Ni-agarose column. The concentration of the purified protein was determined by its predicted extinction coefficient (81,820 M⁻¹ cm⁻¹ at 280 nm). The active CrpE concentration was determined using previously described methods (51).

Generation of Cryptophycin 4 Using the Cryptophycin TE. Synthesis of seco-SNAC-cryptophycin 4 and heterologous expression and purification of the Crp TE (CrpD TE) were conducted following published procedures (16). Seco-SNAC-cryptophycin 4 (2.4 mg) was dissolved to 2 mM in DMSO and then diluted to 100 μ M in 0.1 M Tris pH 7.0. A 10 μ M portion of CrpD TE was added to this mixture and incubated at 30 °C for 15 h. The reaction was next extracted with ethyl acetate (3 \times 25 mL), and the ethyl acetate fractions were pooled and concentrated. The total contents of the ethyl acetate extraction were separated by semi-preparative reversed-phase HPLC (C18 Econosil, 10 \times 250 mm, 5 mL min⁻¹, 10–100% acetonitrile/water, 50 min). The peak corresponding to cryptophycin 4 was collected and lyophilized. HPLC t_{ret} = 36.3 min; mass spectrometry (MS) (ESI⁺) m/z 605.2, [M + H]⁺ (C₃₅H₄₅N₂O₇ requires 605.3).

CrpE Reactions and HPLC Analysis. CrpE reactions contained 100 μ g mL⁻¹ ferredoxin, 0.2 units mL⁻¹ of ferredoxin-NADP+ reductase, 1.4 mM of NADPH, 10 mM of glucose-6-phosphate, 8 units mL⁻¹ of glucose-6-phosphate dehydrogenase, 5.2 μ M seco-SNAC-cryptophycin 4, and 1.8 μ M CrpD TE or 9 μ M cryptophycin 4 in 100 μ L of lysis buffer. The reaction mixtures were incubated at 30 °C for 2 min, and then 3 μ M CrpE and/or 1.8 μ M CrpD TE was added to initiate the reaction at 30 °C for 2–4 h. A control reaction was run in parallel that contained boiled CrpE in place of active CrpE. The reaction mixtures were then separated by analytical reversed-phase HPLC (C18 Econosil, 4.6 \times 250 mm, 1 mL min⁻¹, 30–100% acetonitrile/water + 0.1% TFA, 40 min, 218 nm). Reaction of CrpE with cryptophycin 4 yielded cryptophycin 2 as determined by LC/MS t_{ret} = 25.3 min; MS (ESI⁺) m/z 621.2, [M + H]⁺ (C₃₅H₄₅N₂O₈ requires 621.3), and m/z 643.2, [M + Na]⁺ (C₃₅H₄₄N₂NaO₈ requires 643.3). A large-scale reaction was performed using the same ratio of reagents but 100 \times the volume. For this reaction, spinach ferredoxin reductase was purified from spinach leaves using modified protocols (52). Briefly, 3 kg of crude homogenate obtained from grinding in a Waring blender was fractionated with acetone according to a previously published method (53). The precipitate obtained was dissolved in 58 mM Tris, pH 7.5, and passed through three Hi-Trap Q 5 mL columns in series using an isocratic flow of 5 mL min⁻¹ in the same buffer. The yellow flow-through fractions were concentrated and used directly in the reaction. Purified epoxidized material was analyzed by NMR using a Varian INOVA 600 NMR spectrometer, equipped with a 5 mm HCN probe. The Crp sample was dissolved in 200 μ L of CD₃OD. The NMR spectrum was identical to that reported for cryptophycin 2 with the chemical shifts for the epoxide protons at 3.94 (dd, J = 7.5, 1.8 Hz), 3.70 (d, J = 1.8 Hz), and 1.80 (m) for the adjacent ring proton (42).

General Chemical Procedures: Labeled Precursors, Stable-Isotope Feedings, and Directed Biosynthesis. Sodium [1,2-¹³C₂] acetate, sodium [2-¹³C, ²H₃]acetate, sodium [1-¹³C, ¹⁸O]acetate, sodium [U-¹³C₃]pyruvate, L-[methyl-¹³C]methionine (¹³C, 96%), L-[U-¹³C₉, ¹⁵N]phenylalanine, L-[²H₈]phenylalanine, [1-¹³C]phenylacetic acid, *p*-tolylacetic acid, L-[1-¹³C]tyrosine, DL-[2-¹³C, ¹⁵N] aspartic acid, DL-[2,3,3-²H₃]aspartic acid, L-[1-¹³C]leucine, and

[1-¹³C]2-ketocaproic acid were obtained from the Aldrich Chemical Co. The methods for all of the feeding study experiments are contained within the Supporting Information section. In general, *Nostoc* sp. GSV 224 or *Nostoc* sp. ATCC 53789 were cultured in 20-L glass carboys as previously described (3). All feeding studies were performed with both *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV 224 with essentially identical results. The isotopically labeled or biosynthetic precursor was added to each of two 20-L cultures of *Nostoc* sp. GSV 224 or *Nostoc* sp. ATCC 53789 beginning on day 10 after inoculation unless noted otherwise. For precursor-directed biosynthesis, an aqueous solution of the precursor was added all at once or in aliquots (2–8) every other day. Usually an amino acid precursor was dissolved in 0.5 N HCl to a concentration of 10–30 mg mL⁻¹ and fed in 8 aliquots. Typically, a 0.5 mL portion of the solution was added to each of the 2–4 carboys of *Nostoc* sp. GSV 224 culture in 2-d intervals beginning on day 10–12 after inoculation. After 6–10 additions of the amino acid solution, the cultures were allowed to grow for an additional 3–5 d and then harvested. ¹H and ¹³C NMR spectra were obtained at 500 and 125 MHz in CDCl₃. All ¹³C, ¹⁵N, ²H, and ¹⁸O-labeled compounds used in this study were 99, 96–99, 98, and 95 atom %, respectively, unless noted otherwise. Low-resolution MS was performed at the University of Michigan Mass Spectrometry Laboratory on a Waters Ultima magnetic sector mass spectrometer equipped with an electrospray interface.

Isolation of Labeled Cryptophycins. Lyophilized *Nostoc* sp. GSV 224 (10–30 g) was extracted with 4:1 mixture of CH₃CN/CH₂Cl₂ (40 mL g⁻¹) for 48 h and the extract concentrated *in vacuo* to give a dark green solid. This solid (1 g) was applied to an ODS-coated silica column (55 g, 15 × 2.5 cm) and subjected to flash chromatography with 1:3 CH₃CN/H₂O (0.4 L), 1:1 CH₃CN/H₂O (0.4 L), 65:35 CH₃CN/H₂O (0.8 L), CH₃OH (0.4 L), and CH₂Cl₂ (0.4 L). The material was eluted with 65:35 CH₃CN/H₂O (100–400 mg) and further separated by reversed-phase HPLC (Econosil C18, 10 μm, 25 cm × 22 mm, UV detection at 254 nm, 65:35 CH₃CN/H₂O, flow rate 6 mL min⁻¹) to give labeled cryptophycin 1 (*t*_R = 53.0 min) and a number of fractions containing mixtures of other labeled cryptophycins. The natural cryptophycins eluted between 25 and 100 min. Labeled analogues were isolated using previously described procedures (3). The isolation of deuterated cryptophycins was monitored by ²H NMR spectroscopy. The methods and results of the feeding experiments, including all associated structure elucidation information are contained within the Supporting Information section available online.

Acknowledgments: We kindly acknowledge Zhaohui Xu for pSj8 used for CrpE expression, Sabine Grischow for LC/MS analysis, and Frank Schroeder for NMR analysis. We thank Wu Du, Ghosen Ye, and Jian Liang for expert technical assistance. This research was supported by National Institutes of Health (NIH) Postdoctoral Training Fellowship (NCI CA09676) to Z.Q.B., NIH Grants CA83155 and CA108874 and the Searle Professorship to D.H.S., and NIH Grant CA12623 and National Science Foundation Grant CHE-9530794 to R.E.M.

Supporting Information Available: This material is free of charge via the Internet.

REFERENCES

- Chaganty, S.; Golakoti, T.; Heltzel, C.; Moore, R. E.; and Yoshida, W. Y. (2004) Isolation and structure determination of cryptophycins 38, 326, and 327 from the terrestrial cyanobacterium *Nostoc* sp. GSV 224, *J. Nat. Prod.* 67, 1403–1406.
- Schwartz, R. E.; Hirsch, C. F.; Sesin, D. F.; Flor, J. E.; Chartrain, M.; Fromtling, R. E.; Harris, G. H.; Salvatore, M. J.; Liesch, J. M.; and Yudin, K. (1990) Pharmaceuticals from cultured algae, *J. Ind. Microbiol. Biotechnol.* 5, 113–123.
- Golakoti, T.; Ohtani, I.; Patterson, G. M. L.; Moore, R. E.; Corbett, T. H.; Valeriote, F. A.; and Demchik, L. (1994) Total structures of cryptophycins, potent antitumor depsipeptides from the blue-green-alga *Nostoc* sp. strain GSV-224, *J. Am. Chem. Soc.* 116, 4729–4737.
- Subbaraju, G. V.; Golakoti, T.; Patterson, G. M.; and Moore, R. E. (1997) Three new cryptophycins from *Nostoc* sp. GSV 224, *J. Nat. Prod.* 60, 302–305.
- Golakoti, T.; Yoshida, W. Y.; Chaganty, S.; and Moore, R. E. (2001) Isolation and structure determination of nostocyclopeptides A1 and A2 from the terrestrial cyanobacterium *Nostoc* sp. ATCC 53789, *J. Nat. Prod.* 64, 54–59.
- Smith, C. D.; Zhang, X.; Mooberry, S. L.; Patterson, G. M.; and Moore, R. E. (1994) Cryptophycin: a new antimicrotubule agent active against drug-resistant cells, *Cancer Res.* 54, 3779–3784.
- Lu, K.; Dempsey, J.; Schultz, R. M.; Shih, C.; and Teicher, B. A. (2001) Cryptophycin-induced hyperphosphorylation of Bcl-2, cell cycle arrest and growth inhibition in human H460 NSCLC cells, *Cancer Chemother. Pharmacol.* 47, 170–178.
- Edelman, M. J.; Gandara, D. R.; Hausner, P.; Israel, V.; Thornton, D.; DeSanto, J.; and Doyle, L. A. (2003) Phase 2 study of cryptophycin 52 (LY355703) in patients previously treated with platinum based chemotherapy for advanced non-small cell lung cancer, *Lung Cancer* 39, 197–199.
- D'Agostino, G.; del Campo, J.; Mellado, B.; Izquierdo, M. A.; Minarik, T.; Cirri, L.; Marini, L.; Perez-Gracia, J. L.; and Scambia, G. (2006) A multicenter phase II study of the cryptophycin analog LY355703 in patients with platinum-resistant ovarian cancer, *Int. J. Gynecol. Cancer* 16, 71–76.
- Liang, J.; Moore, R. E.; Moher, E. D.; Munroe, J. E.; Al-awar, R. S.; Hay, D. A.; Varie, D. L.; Zhang, T. Y.; Aikins, J. A.; Martinelli, M. J.; Shih, C.; Ray, J. E.; Gibson, L. L.; Vasudevan, V.; Polin, L.; White, K.; Kushner, J.; Simpson, C.; Pugh, S.; and Corbett, T. H. (2005) Cryptophycins-309, 249 and other cryptophycin analogs: preclinical efficacy studies with mouse and human tumors, *Invest. New Drugs* 23, 213–224.
- Eggen, M.; and Georg, G. I. (2002) The cryptophycins: their synthesis and anticancer activity, *Med. Res. Rev.* 22, 85–101.
- Barrow, R. A.; Hemscheidt, T.; Liang, J.; Paik, S.; Moore, R. E.; and Tius, M. A. (1995) Total synthesis of cryptophycins—revision of the structures of cryptophycin-A and cryptophycin-C, *J. Am. Chem. Soc.* 117, 2479–2490.
- Gardinier, K. M.; and Leahy, J. W. (1997) Enantiospecific total synthesis of the potent antitumor macrolides cryptophycins 1 and 8, *J. Org. Chem.* 62, 7098–7099.
- Chang, H. T.; and Sharpless, K. B. (1996) Molar scale synthesis of enantiopure stilbene oxide, *J. Org. Chem.* 61, 6456–6457.
- Liang, J.; Moher, E. D.; Moore, R. E.; and Hoard, D. W. (2000) Synthesis of cryptophycin 52 using the Sharpless asymmetric dihydroxylation: diol to epoxide transformation optimized for a base-sensitive substrate, *J. Org. Chem.* 65, 3143–3147.
- Beck, Z. Q.; Aldrich, C. C.; Magarvey, N. A.; Georg, G. I.; and Sherman, D. H. (2005) Chemoenzymatic synthesis of cryptophycin/arenastatin natural products, *Biochemistry* 44, 13457–13466.
- Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V. A.; Willis, C. L.; and Sherman, D. H. (2002) The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit, *Gene* 296, 235–247.
- Chang, Z.; Sitachitta, N.; Rossi, J. V.; Roberts, M. A.; Flatt, P. M.; Jia, J.; Sherman, D. H.; and Gerwick, W. H. (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*, *J. Nat. Prod.* 67, 1356–1367.
- Meeks, J. C.; Elhai, J.; Thiel, T.; Potts, M.; Larimer, F.; Lamerdin, J.; Predki, P.; and Atlas, R. (2001) An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium, *Photosyn. Res.* 70, 85–106.

20. Hunsucker, S. W., Klage, K., Slaughter, S. M., Potts, M., and Helm, R. F. (2004) A preliminary investigation of the *Nostoc punctiforme* proteome, *Biochem. Biophys. Res. Commun.* 317, 1121–1127.
21. Challis, G. L., Ravel, J., and Townsend, C. A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, *Chem. Biol.* 7, 211–224.
22. Edwards, D. J., Marquez, B. L., Nogle, L. M., McPhail, K., Goeger, D. E., Roberts, M. A., and Gerwick, W. H. (2004) Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*, *Chem. Biol.* 11, 817–833.
23. Hicks, L. M., Moffitt, M. C., Beer, L. L., Moore, B. S., and Kelleher, N. L. (2006) Structural characterization of *in vitro* and *in vivo* intermediates on the loading module of microcystin synthetase, *ACS Chem. Biol.* 1, 93–102.
24. Caffrey, P. (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases, *Chembiochem* 4, 654–657.
25. Reid, R., Piagentini, M., Rodríguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C. R., and McDaniel, R. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases, *Biochemistry* 42, 72–79.
26. Siskos, A. P., Baerga-Ortiz, A., Bali, S., Stein, V., Mamdani, H., Spiteller, D., Popovic, B., Spencer, J. B., Staunton, J., Weissman, K. J., and Leadlay, P. F. (2005) Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases, *Chem. Biol.* 12, 1145–1153.
27. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis, *Chem. Rev.* 97, 2651–2673.
28. Rouhiainen, L., Paulin, L., Suomalainen, S., Hyttiäinen, H., Buikema, W., Haselkom, R., and Sivonen, K. (2000) Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90, *Mol. Microbiol.* 37, 156–167.
29. Yeh, E., Gameau, S., and Walsh, C. T. (2005) Robust *in vitro* activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebecamycin biosynthesis, *Proc. Natl. Acad. Sci. U.S.A.* 102, 3960–3965.
30. Dorrestein, P. C., Yeh, E., Gameau-Tsodikova, S., Kelleher, N. L., and Walsh, C. T. (2005) Dichlorination of a pyrrolyl-S-carrier protein by FADH₂-dependent halogenase PItA during pyoluteorin biosynthesis, *Proc. Natl. Acad. Sci. U.S.A.* 102, 13843–13848.
31. Jackowski, S. (1996) Biosynthesis of pantothenic acid and coenzyme A, in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., Ed.), pp 687–694, American Society for Microbiology, Washington, DC.
32. Wolf, J. P., and Rapoport, H. (1989) Conformationally constrained peptides. Chiro-specific synthesis of 4-alkyl-substituted gamma-lactam-bridged dipeptides from L-aspartic acid, *J. Org. Chem.* 54, 3164–3173.
33. Ogasawara, Y., Katayama, K., Minami, A., Otsuka, M., Eguchi, T., and Kakinuma, K. (2004) Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in *Streptomyces halstedii*, *Chem. Biol.* 11, 79–86.
34. Nishida, H., Eguchi, T., and Kakinuma, K. (2001) Amino acid starter unit in the biosynthesis of macrolactam polyketide antitumor antibiotic vicenistatin, *Tetrahedron* 57, 8237–8242.
35. Umbarger, H. E. (1996) Biosynthesis of branched-chain amino acids, in *Escherichia coli and Salmonella: Cellular and Molecular Biology* 1 (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) 2nd ed., pp 442–457, American Society for Microbiology, Washington, DC.
36. Cheng, Y. Q. (2006) Deciphering the biosynthetic codes for the potent anti-SARS-CoV cyclodepsipeptide valinomycin in *Streptomyces tsusimaensis* ATCC 15141, *Chembiochem* 7, 471–477.
37. Ehling-Schulz, M., Fricker, M., Grallert, H., Rieck, P., Wagner, M., and Scherer, S. (2006) Cereulide synthetase gene cluster from emetic *Bacillus cereus*: structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1, *BMC Microbiol.* 6, 20.
38. Magarvey, N. A., Ehling-Schulz, M., and Walsh, C. T. (2006) Characterization of the cereulide NRPS α -hydroxy acid specifying modules: activation of α -keto acids and chiral reduction on the assembly line, *J. Am. Chem. Soc.* 128, 10698–10699.
39. Zaleta-Rivera, K., Xu, C., Yu, F., Butchko, R. A., Proctor, R. H., Hidalgo-Lara, M. E., Raza, A., Dussault, P. H., and Du, L. (2006) A bidomain nonribosomal peptide synthetase encoded by FUM14 catalyzes the formation of tricarballic esters in the biosynthesis of fumonisins, *Biochemistry* 45, 2561–2569.
40. Pee, K.-H. v., and Patallo, E. P. (2006) Flavin-dependent halogenases involved in secondary metabolism in bacteria, *Appl. Microbiol. Biotechnol.* 70, 631–641.
41. Al-Awar, R. S., Ray, J. E., Schultz, R. M., Andis, S. L., Kennedy, J. H., Moore, R. E., Liang, J., Golakoti, T., Subbaraju, G. V., and Corbett, T. H. (2003) A convergent approach to cryptophycin 52 analogues: synthesis and biological evaluation of a novel series of fragment A epoxides and chlorohydrins, *J. Med. Chem.* 46, 2985–3007.
42. Ghosh, A. K., and Bischoff, A. (2000) A convergent synthesis of (+)-cryptophycin B, a potent antitumor macrolide from *Nostoc* sp. cyanobacteria, *Org. Lett.* 2, 1573–1575.
43. Rikkinen, J., Oksanen, I., and Lohtander, K. (2002) Lichen guilds share related cyanobacterial symbionts, *Science* 297, 357.
44. Golden, S. S., Brusslan, J., and Haselkom, R. (1987) Genetic engineering of the cyanobacterial chromosome, *Methods Enzymol.* 153, 215–231.
45. Sambrook, J., and Russell, D. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
46. Moffitt, M. C., and Neilan, B. A. (2004) Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins, *Appl. Environ. Microbiol.* 70, 6353–6362.
47. Neilan, B. A., Dittmann, E., Rouhiainen, L., Bass, R. A., Schaub, V., Sivonen, K., and Borner, T. (1999) Nonribosomal peptide synthesis and toxicity of cyanobacteria, *J. Bacteriol.* 181, 4089–4097.
48. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* 25, 4876–4882.
49. Cohen, M. F., Wallis, J. G., Campbell, E. L., and Meeks, J. C. (1994) Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives, *Microbiology* 140, 3233–3240.
50. Ansari, M. Z., Yadav, G., Gokhale, R. S., and Mohanty, D. (2004) NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases, *Nucleic Acids Res.* 32, W405–W413.
51. Omura, T., and Sato, R. (1964) The carbon monoxide-binding pigment of liver microsomes. II. Solubilization, purification, and properties, *J. Biol. Chem.* 239, 2379–2385.
52. Schurmann, P. (1995) Ferredoxin: thioredoxin system, *Methods Enzymol.* 252, 274–283.
53. Shin, M. (1971) Ferredoxin-NADP reductase from spinach, *Methods Enzymol.* 23, 440–447.
54. Donadio, S., and Katz, L. (1992) Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*, *Gene* 111, 51–60.