Biosynthesis and Structures of Cyclomarins and Cyclomarazines, Prenylated Cyclic Peptides of Marine Actinobacterial Origin

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · MAY 2008

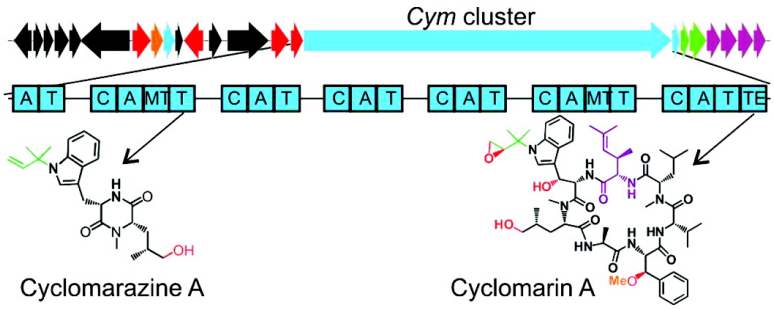
Impact Factor: 12.11 · DOI: 10.1021/ja711188x · Source: PubMed

Article

Biosynthesis and Structures of Cyclomarins and Cyclomarazines, Prenylated Cyclic Peptides of Marine Actinobacterial Origin

Andrew W. Schultz, Dong-Chan Oh, John R. Carney, R. Thomas Williamson, Daniel W. Udwary, Paul R. Jensen, Steven J. Gould, William Fenical, and Bradley S. Moore

J. Am. Chem. Soc., **2008**, 130 (13), 4507-4516• DOI: 10.1021/ja711188x • Publication Date (Web): 11 March 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 1 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 03/11/2008

Biosynthesis and Structures of Cyclomarins and Cyclomarazines, Prenylated Cyclic Peptides of Marine Actinobacterial Origin

Andrew W. Schultz,[†] Dong-Chan Oh,[†] John R. Carney,[‡] R. Thomas Williamson,[‡] Daniel W. Udwary,[†] Paul R. Jensen,[†] Steven J. Gould,[‡] William Fenical,[†] and Bradley S. Moore*[,†,§]

Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0204, Department of Chemistry, Oregon State University, Corvallis, Oregon 97331-4003, and The Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093

Received December 17, 2007; E-mail: bsmoore@ucsd.edu

Abstract: Two new diketopiperazine dipeptides, cyclomarazines A and B, were isolated and characterized along with the new cyclic heptapeptide cyclomarin D from the marine bacterium *Salinispora arenicola* CNS-205. These structurally related cyclic peptides each contain modified amino acid residues, including derivatives of *N*-(1,1-dimethylallyl)-tryptophan and δ-hydroxyleucine, which are common in the di- and heptapeptide series. Stable isotope incorporation studies in *Streptomyces* sp. CNB-982, which was first reported to produce the cyclomarin anti-inflammatory agents, illuminated the biosynthetic building blocks associated with the major metabolite cyclomarin A, signifying that this marine microbial peptide is nonribosomally derived largely from nonproteinogenic amino acid residues. DNA sequence analysis of the 5.8 Mb *S. arenicola* circular genome and PCR-targeted gene inactivation experiments identified the 47 kb cyclomarin/cyclomarazine biosynthetic gene cluster (*cym*) harboring 23 open reading frames. The *cym* locus is dominated by the 23 358 bp *cymA*, which encodes a 7-module nonribosomal peptide synthetase (NRPS) responsible for assembly of the full-length cyclomarin heptapeptides as well as the truncated cyclomarazine dipeptides. The unprecedented biosynthetic feature of the megasynthetase CymA to synthesize differently sized peptides in vivo may be triggered by the level of β oxidation of the priming tryptophan residue, which is oxidized in the cyclomarin series and unoxidized in the cyclomarazines. Biosynthesis of the *N*-(1,1-dimethyl-2,3-epoxypropyl)-β-hydroxytryptophan residue of cyclomarin A was further illuminated through gene inactivation experiments, which suggest that the tryptophan residue is reverse prenylated by CymD prior to release of the cyclic peptide from the CymA megasynthetase, whereas the cytochrome P450 CymV installs the epoxide group on the isoprene of cyclomarin C post-NRPS assembly. Last, the novel amino acid residue 2-amino-3,5-dimethylhex-4-enoic acid in the cyclomarin series was shown by bioinformatics and stable isotope experiments to derive from a new pathway involving condensation of isobutyraldehyde and pyruvate followed by *S*-adenosylmethionine methylation. Assembly of this unsaturated, branched amino acid is unexpectedly related to the degradation of the environmental pollutant 3-(3-hydroxyphenyl)propionic acid.

Introduction

Cyclomarins A−C (**1**−**3**) are potent anti-inflammatory marine natural products first described from an estuarine streptomycete, strain CNB-982 (Figure 1).[1] Slight structural variations in oxidation and methylation differentiate these 21-membered cyclic heptapeptides, which are comprised of a number of modified and nonproteinogenic amino acid residues, including *N*-(1,1-dimethyl-2,3-epoxypropyl)-β-hydroxytryptophan, *N*-methyl-δ-hydroxyleucine, β-methoxyphenylalanine, *N*-methylleucine, and 2-amino-3,5-dimethylhex-4-enoic acid (ADH) in addition to unmodified alanine and valine moieties. The syntheses of these unusual residues and the total synthesis of **3** have been described.[2]

ADH is a novel structural unit of **1**−**3** and lacks a clear biosynthetic origin. Although the structure of this eight-carbon, nonproteinogenic amino acid residue is characteristic of proteinogenic branched chain amino acids, its unsaturated, branched side chain has not been encountered in other natural products. Cyclomarin's δ-hydroxyleucine and reverse *N*-prenylated β-hydroxytryptophan units are rare peptidic building blocks with

[†] Scripps Institution of Oceanography, University of California, San Diego.
[‡] Oregon State University.
[§] The Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego.

(1) Renner, M. K.; Shen, Y. C.; Cheng, X. C.; Jensen, P. R.; Frankmoelle, W.; Kauffman, C. A.; Fenical, W.; Lobkovsky, E.; Clardy, J. *J. Am. Chem. Soc.* **1999**, *121*, 11273−11276.

(2) Hamada, Y.; Shioiri, T. *Chem. Rev.* **2005**, *105*, 4441−4482.

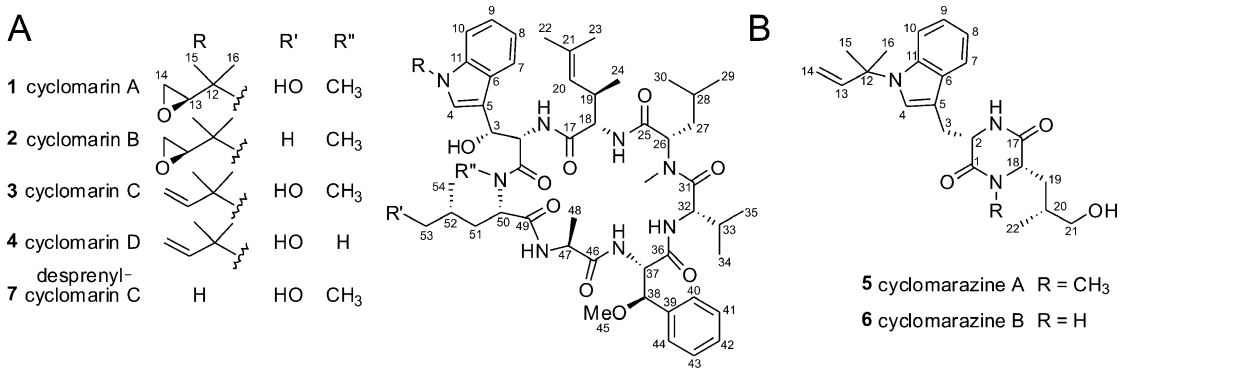


Figure 1. Structures of cyclomarin (**1**−**4**, **7**) and cyclomarazine (**5** and **6**) cyclic peptides from *Salinispora arenicola* CNS-205 and *Streptomyces* sp. CNB-982.

limited distribution. δ-Hydroxyleucine was previously described in depsipeptides from *Paecilomyces lilacinus*[3] and *Biploris zeicola*[4] fungal metabolites where preliminary enzymological studies suggest Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes catalyze hydroxylation at aliphatic carbons of amino acid residues.[5] The reverse *N*-prenylated tryptophan moiety was observed only once before in the ilamycins from *Streptomyces islandicus*.[6] While the prenylation of indoles is commonplace in fungal metabolites such as the ergot alkaloids ergotamine and fumigaclavine, the diketopiperazines fumitremorgin B and brevianamide F and the indole diterpenes paxilline and lolitrem B,[7] the *N*-prenylation of the indole nitrogen of tryptophan is a rare biosynthetic feature in fungi and bacteria. One of the few examples of *N*-prenylated tryptophan derivatives, in addition to the cyclomarins and ilamycins, is flustramine A from the marine bryozoan *Flustra foliacea*.[8] Several enzymes responsible for the transfer of an isoprene to an indole ring have been characterized from bacteria and fungi, including LtxC from the cyanobacterium *Lyngbya majuscula*, which catalyzes the transfer of geranyl pyrophosphate to indolactam-V during the biosynthesis of lyngbyatoxin.[9] Bioinformatic analysis of the genome sequence of *Aspergillus fumigatus* revealed several prenyltransferases (PTases),[10] four of which have been characterized. Three of these enzymes, FtmPT1,[11] FgaPT2,[12] and 7-DMATS,[13] prenylate the indole ring at C2, C4, and C7, respectively, while the other example, CdpNPT, catalyzes a "forward" prenyl transfer to the indole nitrogen.[14] To the best of our knowledge, no enzyme has been characterized that catalyzes the reverse prenylation of an indole nitrogen.

On the basis of these unusual structural features, we set out to explore the biosynthesis of the cyclomarin peptides in *Streptomyces* sp. CNB-982 with stable isotopes. During the course of this study, we identified new cyclomarin derivatives in another marine actinomycete, *Salinispora arenicola* CNS-205,[15,16] along with the truncated diketopiperazine (DKP) cyclomarazines A−B, which appeared to share a common biogenesis. Sequence analysis of the finished genome of *S. arenicola* CNS-205 (GenBank accession number NC_009953) provided the cyclomarin/cyclomarazine nonribosomal peptide synthetase (NRPS) biosynthetic gene cluster, which furthermore permitted us to explore the biosynthesis of these cyclic peptides at the biochemical and genetic levels. Here we report a comprehensive investigation of cyclomarin biosynthesis involving the description of new structural analogs, precursor incorporation studies, gene cluster sequence analysis, and in vivo functional analysis.

Results

Discovery of New Cyclomarin Derivatives in *S. arenicola*. *Salinispora arenicola* CNS-205 was isolated from Palau[17] and produces potent secondary metabolites belonging to the rifamycin and staurosporin structural families in addition to the known anti-inflammatory cyclic peptides cyclomarin A (**1**) and C (**3**).[18] Further inspection of the culture broth of this marine isolate revealed new cyclomarin analogues (Figure 1), thereby extending the structural diversity of this peptide natural product family that was first discovered in the marine-derived streptomycete CNB-982.[1]

Inspection of 1D and 2D NMR spectral data quickly revealed that cyclomarin D (**4**) differed from **3** by the absence of the *N*-methyl group at C56, which was further supported by HR-ESI mass spectrometry to give the molecular formula $C_{55}H_{80}N_8O_{10}$. The diketopiperazine cyclomarazine A (**5**), on the other hand, was considerably smaller at $C_{23}H_{31}N_3O_3$ and shared two amino acid residues with that of the cyclic heptapeptide **3**. Interpretation of 2D NMR spectra (gCOSY, gHSQC, and gHMBC) elucidated *N*-(1,1-dimethyl-allyl)-tryptophan and *N*-methyl-δ-hydroxyleucine residues, and HMBC correlations from 18-NMe and H2 to C1 and from 2-NH and H18 to C17 built the diketopiperazine ring system. The absolute configurations of the amino acid α-carbons were determined by the advanced Marfey method,[19,20] while the configuration at C20

(3) Isogai, A.; Nakayama, J.; Takayama, S.; Kusai, A.; Suzuki, A. *Biosci. Biotechnol. Biochem.* **1992**, *56*, 1079−1085.
(4) Ueda, K.; Xiao, J. Z.; Doke, N.; Nakatsuka, S. *Tetrahedron Lett.* **1992**, *33*, 5377−5380.
(5) Hausinger, R. P. *Rev. Biochem. Mol. Biol.* **2004**, *39*, 21−68.
(6) Takita, T.; Ohi, K.; Okami, Y.; Maeda, K.; Umezawa, H. *J. Antibiot.* **1962**, *15*, 46−48.
(7) Hoffmeister, D.; Keller, N. P. *Nat. Prod. Rep.* **2007**, *24*, 393−416.
(8) Carle, J. S.; Christophersen, C. *J. Org. Chem.* **1980**, *45*, 1586−1589.
(9) Edwards, D. J.; Gerwick, W. H. *J. Am. Chem. Soc.* **2004**, *126*, 11432−11433.
(10) Nierman, W. C. et al. *Nature* **2005**, *438*, 1151−1156.
(11) Grundmann, A.; Li, S.-M. *Microbiology* **2005**, *151*, 2199−2207.
(12) Unsold, I. A.; Li, S.-M. *Microbiology* **2005**, *151*, 1499−1505.
(13) Kremer, A.; Westrich, L.; Li, S.-M. *Microbiology* **2007**, *153*, 3409−3416.
(14) Yin, W.-B.; Ruan, H.-L.; Westrich, L.; Grundmann, A.; Li, S.-M. *ChemBioChem* **2007**, *8*, 1154−1161.
(15) Maldonado, L.; Fenical, W.; Goodfellow, M.; Jensen, P. R.; Ward, A. C. *Int. J. System. Appl. Microbiol.* **2005**, *55*, 1759−1766.
(16) Jensen, P. R.; Williams, P. G.; Oh, D.-C.; Zeigler, L.; Fenical, W. *Appl. Environ. Microbiol.* **2007**, *73*, 1146−1152.
(17) Gontang, E. A.; Fenical, W.; Jensen, P. R. *Appl. Environ. Microbiol.* **2007**, *73*, 3272−3282.
(18) Fenical, W.; Jensen, P. R. *Nat. Chem. Biol.* **2006**, *2*, 666−673.
(19) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 3346−3352.
(20) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146−5151.

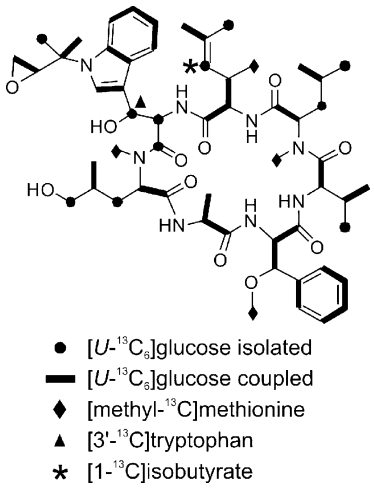


Figure 2. Biosynthetic origin of the carbons in **1**. ^{13}C-Labeling patterns from the incorporation of [U-$^{13}C_6$]glucose (bold lines and dots), [methyl-^{13}C]methionine (diamonds), [3′-^{13}C]tryptophan (triangle), and [1-^{13}C]-isobutyrate (asterisk) are depicted. Intact incorporation of $^{13}C-^{13}C$ units was determined by INADEQUATE NMR.

was established by consecutive *J*-based configuration analysis[21] from C18 to C20. These analyses revealed the same absolute configurations as in the corresponding residues in **3**. Spectral analysis of the related metabolite cyclomarazine B (**6**) indicated one additional exchangeable proton and the lack of the *N*-methyl group, which was consistent with a DKP composed of *N*-(1,1-dimethyl-allyl)-tryptophan and δ-hydroxyleucine residues. Cyclomarazine A and B displayed identical Cotton effects in the CD spectra with a positive effect at 232−234 nm and a negative effect at 216−218 nm, supporting the same absolute configuration. A key structural difference between the heptapeptides **1**−**4** and the DKPs **5** and **6** involves the β-position of the tryptophan unit, which is hydroxylated only in the larger cyclic peptides **1**−**4**.

Cyclomarazines A and B were identified as moderate antimicrobial agents possessing MIC values of 18 and 13 μg/mL against methicillin-resistant *Staphylococcus areus* and vancomycin-resistant *Entrococcus faecium*, respectively. They had no significant antifungal activity against amphotericin-resistant *Candida albicans* or cytotoxicity against the human colon carcinoma cell line HCT-116. Cyclomarin D displayed modest HCT-116 cytotoxicity with an IC_{50} value of 2 μg/mL and no antifungal or antimicrobial activities.

Molecular Basis for the Biosynthesis of the Cyclomarins. Structural inspection of the cyclomarins and cyclomarazines suggested that these cyclic peptides are biosynthesized nonribosomally[22] since they harbor modified amino acid residues (*N*-methylation, *N*-prenylation, and side chain oxidation) as well as the novel nonproteinogenic amino acid residue 2-amino-3,5-dimethyl-4-hexenoic acid (ADH). Feeding experiments with [U-^{13}C]glucose, [methyl-^{13}C]methionine, and [3′-^{13}C]tryptophan in *Streptomyces* sp. CNB-982 established the biosynthetic origins of the precursor building blocks in **1**, thereby confirming that all of the modified residues except for ADH are derived from standard amino acids (Figure 2, Table 1). Four methyl carbons in **1**, namely, the two *N*-methyl amides, the *O*-methyl group at C38, and the C24 methyl of the ADH residue, are derived from methionine. Taken together, the biosynthesis of **1** is expected to involve a seven-module NRPS that assembles largely modified amino acid residues involving four methyltransferases (MTs), four oxygenases, a prenyltransferase (PTase), and a dedicated pathway to ADH.

We recently completed the genome sequence analysis of the cyclomarin/cyclomarazine-producing bacterium *S. arenicola* CNS-205, which revealed 10 NRPS-related biosynthetic gene clusters in its 5 786 361 bp genome (Udwary, D. W.;Penn, K.;

Table 1. NMR Data for [U-^{13}C]Glucose-Derived Cyclomarin A (**1**)

residue	C no.	^{13}C	$^{1}J_{cc}$ (Hz)	INADEQUATE
N-(1,1-dimethyl-2,3-epoxypropyl)-β-OH-Trp	1	171.0		
	2	52.9		
	3	68.6		
	4	123.4	71	C5
	5	111.9	71	C4
	6	127.0	56	C11
	7	119.2	a	
	8	119.8	*a*	
	9	122.1	*a*	
	10	113.7	*a*	
	11	136.1	56	C6
	12	58.1	38	C16, C15[b]
	13	57.8	29	C14
	14	45.4	29	C13
	15	23.1	38	C12[b]
	16	24.5	38	C12
ADH	17	172.5	54	C18
	18	58.1	55	C17, C19[c]
	19	35.5	36	C18[d]
	20	124.8		
	21	134.4	42	C23
	22	25.7		
	23	18.9	42	C21
	24	18.5		
N-Me-Leu	25	168.4	52	C26
	26	58.6	52	C25
	27	38.9		
	28	25.0	35	C29
	29	24.4	35	C28
	30	23.5		
	26-*N*-Me	29.6		
Val	31	170.6	55	C32
	32	55.2	55	C31
	33	30.8	35	C35
	34	19.3		
	35	20.0	35	C33
β-OMe-Phe	36	169.6	50	C37
	37	55.9	50, 40	C36, C38
	38	80.0	40	C37
	39	135.1	57	C44
	40	127.8	57	
	41	128.3	*a*	
	42	128.7	*a*	
	43	128.3	*a*	
	44	127.8	57	
	45	57.8	*a*	
Ala	46	171.6	55	C47
	47	50.6	55, 35	C46, C48
	48	20.8	35	C47
N-Me-δ-OH-Leu	49	168.8	52	C50
	50	59.3	52	C49
	51	33.1		
	52	33.2	35	C54
	53	66.5		
	54	17.6	35	C52
	N-8-Me	29.2		

[a] Low-intensity coupled signals were observed, but ^{1}J was not determined. [b] Minor coupling species indicative of the MEP pathway. [c] Single enrichment also evident. [d] Two coupling species identified: doublet ($J_{C17-C18}$) and doublet of doublets ($J_{C17-C18}/J_{C18-C19}$).

(21) Matsumori, N.; Kaneno, D.; Murata, M.; Nakamura, H.; Tachibana, K. *J. Org. Chem.* **1999**, *64*, 866−876.
(22) Fischbach, M. A.; Walsh, C. T. *Chem. Rev.* **2006**, *106*, 3468−3496.

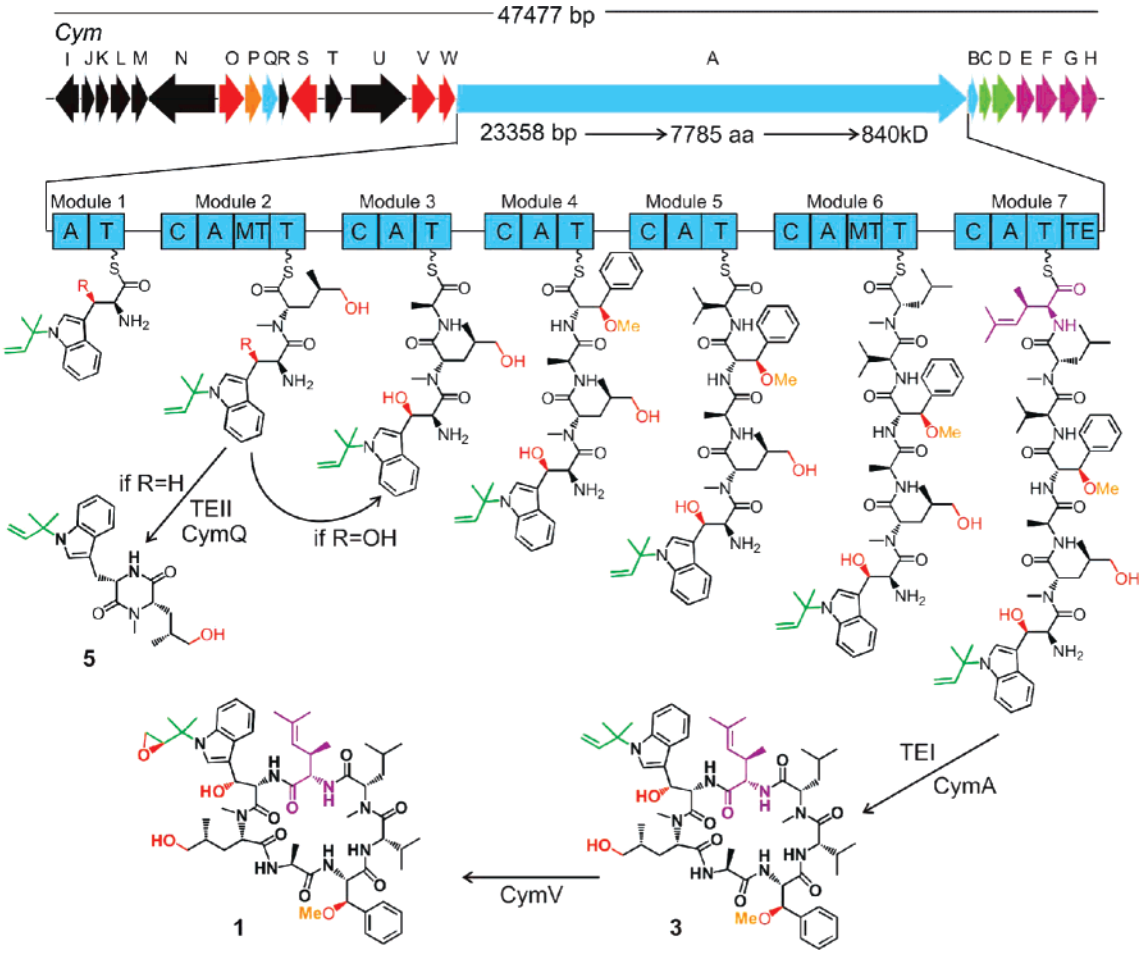


Figure 3. Biosynthetic gene cluster organization of *cym* and proposed biosynthesis of cyclomarin A (**1**) and cyclomarzaine A (**5**). Each arrow represents the direction of transcription of an ORF and is color coded to signify enzyme function which is further reflected chemically. NRPS-related genes are colored blue with enzymatic domain abbreviations as follows: A, adenylation; T, thiolation (peptidyl carrier protein); C, condensation; MT, methyltransferase; and TE, thioesterase. Oxidative genes are colored red, *O*-methyltransferases in orange, prenylation related genes in green, ADH biosynthetic genes in purple, and regulatory/transport/other genes in black.

Jensen, P. R.; Moore, B. S. Unpublished results). Further inspection identified a 47 477 bp biosynthetic gene cluster (*cym*) harboring 23 open reading frames (ORFs) whose deduced functions are consistent with cyclomarin biosynthesis. The *cym* locus is dominated by the largest ORF in the *S. arenicola* genome, the 23 358 bp *cymA*, which codes for a heptamodular NRPS (Figure 3). Its domain architecture is consistent with the cyclomarin heptapeptides in which modules-2 and -6 additionally harbor MT domains, suggesting that the cyclic peptide framework is wholly assembled by this single megasynthetase. Bioinformatic analysis of the seven adenylation (A) domains in CymA strongly correlated module-4 to the activation of L-phenylalanine, while the specificities of the remaining A domains was less clear.[23,24] Taken together, these data suggested the order of incorporation of the amino acid residues in cyclomarin biosynthesis starting with the tryptophan derivative and ending with ADH to yield a linear heptapeptide intermediate bound to the thiolation (T) domain in module-7 (Figure 3). Subsequent release and macrocyclization is then putatively achieved by the C-terminal thioesterase (TE). The cyclomarazine DKPs are consistent with this biosynthetic model in which the module-2-bound diketide is cleaved from CymA to yield the diketopiperazine. This premature cleavage reaction may be facilitated by the type II TE[25] CymQ.

The *cym* cluster further harbors four oxygenases (*cymO*, *cymS*, *cymV*, and *cymW*), a putative PTase (*cymD*), an *O*-MT (*cymP*), a four-gene operon associated with ADH biosynthesis (*cymE*−*H*), four genes putatively involved in regulation and resistance, and two others in phosphoenolpyruvate metabolic flux (Table 2). Pseudogenes containing fragments of integrases and transposases flank the *cym* locus, suggesting that this cluster may have been acquired horizontally. This observation may explain the co-occurrence of this pathway in *S.* sp. CNB-982.

2-Amino-3,5-dimethyl-4-hexenoic Acid (ADH) Biosynthesis. Six of seven amino acid constituents in **1** are clearly recognizable as derived from common protein amino acids modified by methylation, prenylation, and/or hydroxylation. The remaining residue ADH has no proteinogenic counterpart. A number of scenarios involving analogous reactions in branched-chain amino acid biosynthesis were envisaged and first probed in *S.* sp. CNB-982 with stable isotopes. The labeling pattern resulting from [U-^{13}C]glucose and [methyl-^{13}C]methionine assimilation (Figure 2) initially suggested the involvement of leucine, acetate, and methionine building blocks. The C24

(23) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493−505.
(24) Challis, G. L.; Ravel, J.; Townsend, C. A. *Chem. Biol.* **2000**, *7*, 211−224.
(25) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14083−14088.

Table 2. Deduced Functions of the Open Reading Frames in the *cym* Gene Cluster

protein	amino acid	accession	protein family (pfam)	proposed function
CymI	386	SARE4547	phosphoenolpyruvate synthase	PEP synthesis from pyruvate
CymJ	214	SARE4548	no conserved domain	unknown
CymK	200	SARE4549	fructose-2,6-bisphosphatase	glycolysis regulation
CymL	314	SARE4550	ABC-type multidrug transport system	cyclomarin transporter
CymM	264	SARE4551	ABC-type multidrug transport system	cyclomarin transporter
CymN	1043	SARE4552	DnrI, transcriptional activator, SARP	regulation
CymO	404	SARE4553	cytochrome P450	phenylalanine β-hydroxylase
CymP	263	SARE4554	*O*-methyltransferase	*O*-methylation of β-hydroxytryptophan
CymQ	242	SARE4555	type II thioesterase	thioesterase
CymR	163	SARE4556	no conserved domain	unknown
CymS	420	SARE4557	cytochrome P450	leucine δ-hydroxylase
CymT	274	SARE4558	transposase	transposase
CymU	877	SARE4559	LuxR	regulation
CymV	395	SARE4560	cytochrome P450	cyclomarin C epoxidase[a]
CymW	264	SARE4561	dioxygenase	tryptophan β-hydroxylase
CymA	7785	SARE4562	NRPS	cyclomarin synthetase
CymB	71	SARE4563	MbtH-like protein	NRPS associated
CymC	196	SARE4564	type I IPP isomerase	isopentyenyl diphosphate isomerase
CymD	373	SARE4565	no conserved domain[b]	*N*-prenyltransferase[a]
CymE	295	SARE4566	acetaldehyde dehydrogenase	ADH[c] biosynthesis
CymF	348	SARE4567	4-hydroxy-2-ketovalerate aldolase	ADH biosynthesis
CymG	327	SARE4568	*O*-methyltransferase	ADH biosynthesis
CymH	264	SARE4569	2-keto-4-pentenoate hydratase	ADH biosynthesis

[a] Function demonstrated in this work. [b] No entry reported in pfam database for soluble aromatic prenytransferases. [c] 2-Amino-3,5-dimethyl-4-hexenoic acid.

methyl group is clearly derived from methionine, whereas the remaining ADH carbons originate from glucose fragments. The origin of the C17−C19 three-carbon unit in **1** initially led us astray due to two discernible ^{13}C-labeling patterns involving intact and fragmented incorporations of glucose units in which the ^{13}C NMR signal of C19 was singly enriched as well as coupled to C18, which in turn couples with C17. At the time, we incorrectly assumed that the coupling between C18 and C19 was an artifact arising from interunit coupling. Hence, our working hypothesis to ADH biosynthesis involved a pathway by analogy to leucine biosynthesis in which α-ketoisocaproic acid could condense with acetyl-CoA followed by isomerization, oxidative decarboxylation, β-methylation, desaturation, and transamination.

Inspection of the *S. arenicola cym* cluster revealed the four-gene operon *cymE*−*H* downstream of the *cymA* NRPS, which suggested an entirely different pathway to ADH involving isobutyrate, pyruvate, and methionine precursors (Figure 4a). Three of the four genes were homologous to those involved in the degradation of 3-(3-hydroxyphenyl)propionic acid (3-HPP) in *Comamonas testosteroni* TA441.[26] In the catabolism of 3-HPP, the degradant 2-hydroxypenta-2,4-dienoic acid is hydrated by MhpD to 4-hydroxy-2-oxopentanoic acid to facilitate the reverse aldol reaction to pyruvate and acetaldehyde (Figure 4b). Acetaldehyde is then subsequently converted to acetyl-CoA by the NAD$^+$-dependent dehydrogenase MhpF. In an analogous manner yet operating in the reverse biosynthetic direction, valine-derived isobutyraldehyde via isobutyryl-CoA is homologated with pyruvate to give 4-hydroxy-5-methyl-2-oxohexanoic acid via CymE and CymF, respectively (Figure 4a). CymE and CymF share high sequence identity with MhpF (36%; BAA82883) and MhpE (41%; BAA82884), respectively. Dehydration by CymH, which is 35% identical to MhpD (BAA82880), putatively yields 2-hydroxy-5-methylhexa-2,4,-dienoic acid. Methylation by the SAM-dependent MT CymG followed by transamination utilizing a branched-chain amino acid transaminase such as SARE1130 would provide ADH. These last two enzymatic reactions are homologous to those in the biosynthesis of the (2*S*,3*R*)-3-methylglutamate residue of CDA from *Streptomyces coelicolor* A3(2),[27] although CymG shares no significant structural identity with the CDA MT (NP_627429). Support for this revised biosynthetic pathway to ADH was achieved in *S. arenicola* CNS-205 with the administration of sodium [1-^{13}C]isobutyrate, which specifically enriched C20 of **1** at 29%.

Reverse *N*-Prenylation of the Tryptophan Residue. The [U-^{13}C]glucose feeding experiment revealed that the isopentenyl pyrophosphate (IPP)-derived unit attached to the indole nitrogen atom originates in *S.* sp. CNB-982 predominantly from the mevalonate pathway with minor incorporation from the methylerythritol (MEP) pathway (Figure 2). While the presence of both isoprenoid pathways has been demonstrated in other streptomycetes,[28] genomic analysis of *S. arenicola* exclusively revealed the complete MEP pathway. Although the *S. arenicola* MEP pathway genes are dispersed in the genome, the *cym* cluster harbors two terpenoid-related genes, namely, the type I IPP isomerase *cymC* and the PTase *cymD* (Figure 3, Table 2). CymD shares 23% identity to the dimethylallyltyrosine synthase SirD (AAS92554) involved in the biosynthesis of sirodesmin PL from the plant pathogenic fungus *Leptosphaeria maculans*[29] and 24% identity to the PTase LtxC (AAT12285), which is involved in the *C*-geranylation of a tryptophan residue during lyngbyatoxin biosynthesis in the cyanobacterium *Lyngbya majuscula*.[9]

To characterize the function of CymD, we inactivated the encoding gene by PCR-targeted gene replacement utilizing a protocol previously adapted in *Salinispora tropica*.[30] Replace-

(26) Arai, H.; Yamamoto, T.; Ohishi, T.; Shimizu, T.; Nakata, T.; Kudo, T. *Microbiology* **1999**, *145*, 2813−2820.
(27) Mahlert, C.; Kopp, F.; Thirlway, J.; Micklefield, J.; Marahiel, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 12011−12018.
(28) Kuzuyama, T.; Seto, H. *Nat. Prod. Rep.* **2003**, *20*, 171−183.
(29) Gardiner, D. M.; Cozijnsen, A. J.; Wilson, L. M.; Pedras, M. S.; Howlett, B. J. *Mol. Microbiol.* **2004**, *53*, 1307−1318.
(30) Eustáquio, A. S.; Pojer, F.; Noel, J. P.; Moore, B. S. *Nat. Chem. Biol.* **2008**, *4*, 69−74.

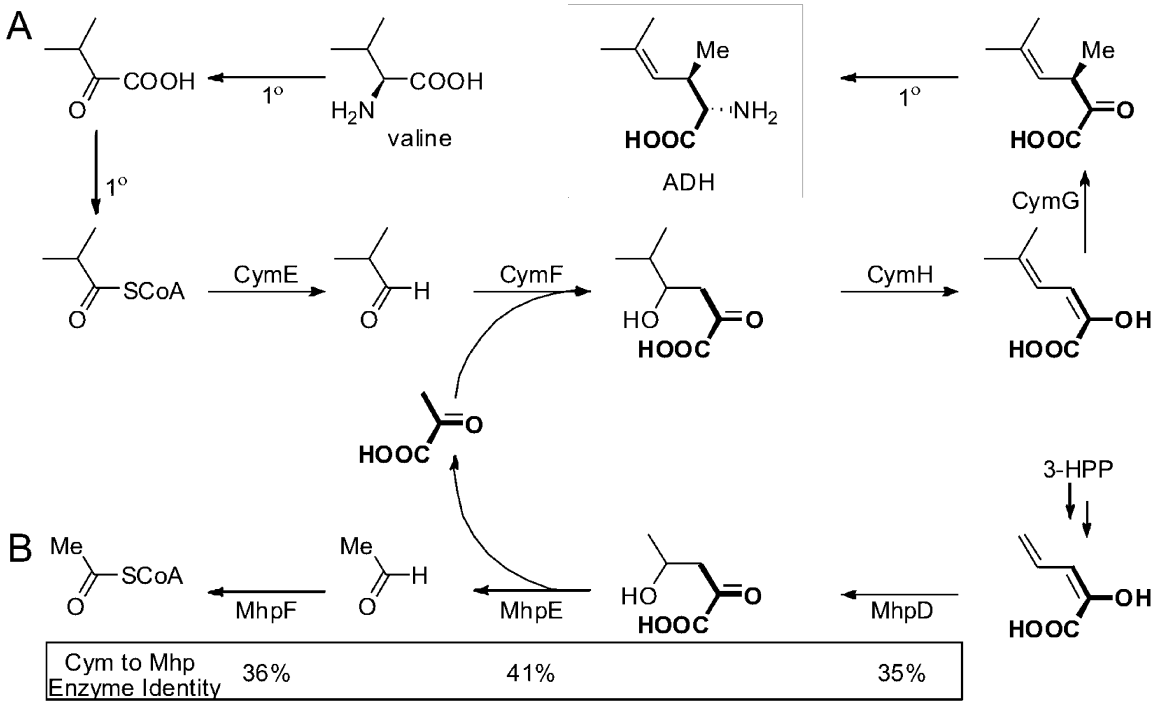


Figure 4. Proposed pathway for the biosynthesis of (A) 2-amino-3,5-dimethyl-4-hexenoic acid (ADH) from valine and pyruvate (bold) in *S. arenicola* CNS-205 and its relatedness to the catabolism of (B) 3-(3-hydroxyphenyl)propionic acid (3-HPP) in *C. testosteroni* TA441 (Mhp pathway). Boxed percentages denote % identity of Cym to corresponding enzymes in the Mhp pathway. 1° denotes primary enzymatic reactions in *S. arenicola* encoded by genes outside the *cym* cluster.

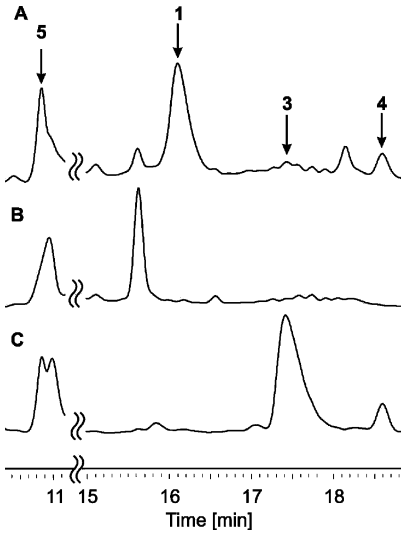


Figure 5. LC-MS analysis of the organic extracts of (A) *S. arenicola* CNS-205, (B) *S. arenicola* CNS-205 *cymD*$^-$, and (C) *S. arenicola* CNS-205 *cymV*$^-$. Chromatograms recorded at 210 nm. Note the complete loss of peaks corresponding to **1** and **3**−**5** in the *cymD* replacement mutant (trace B) and loss of **1** in the *cymV* mutant with a 10-fold concomitant increase in production of **3** (trace C).

ment of *cymD* with an apramycin resistance/*oriT* cassette in the pCC1Fos-based clone BPPW9569 harboring a 38 kb fragment of the *cym* cluster gave the *S. arenicola cymD*$^-$ knockout mutant. HPLC-MS analysis of the mutant verified elimination of cyclomarin and cyclomarazine production (Figure 5). We anticipated the formation of new desprenyl cyclomarin and cyclomarazine analogs in comparable quantities to the wild-type compounds assuming that these molecules would be the substrates for CymD. Although we did isolate and partially characterize desprenylcyclomarin C (**7**) by HR-FTMS (m/z = 981.5408 $[M + Na]^+$, 981.5420 calcd), it was produced in the mutant at approximately 100-fold less than the corresponding **1** in the wild-type strain, thereby impeding its complete chemical characterization.

Oxidative Tailoring. In agreement with the chemical structure of cyclomarin A (**1**), the *cym* gene cluster harbors four oxygenase-encoding genes. BLAST analysis of the Fe(II)/α-ketoglutarate-dependent dioxygenase CymW and the three cytochrome P450s CymO, CymS, and CymV provided insight into their putative biosynthetic functions in amino acid modification (Table 2). While there are no ferrodoxins associated with the *cym* locus to accompany the three P450s, there are eight ferrodoxin encoding genes distributed throughout the *S. arenicola* genome.

CymO belongs to a family of NRPS-associated amino acid β-hydroxylases that oxidize the β-position of PCP-bound amino acid residues.[31] Cyclomarin A harbors two such residues, β-methoxyphenylalanine and β-hydroxytryptophan. Immediately downstream of *cymO* is the putative MT *cymP*, which strongly suggests that their protein products are jointly involved in the β-oxidation/methylation of CymA-(T_4)-bound phenylalanine prior to the condensation with the CymA-module-3 tripeptide precursor (Figure 3). The coding specificity of the module-4 A domain is strongly consistent with an unmodified phenylalanine substrate,[23,24] further supporting the hypothesis that these amino acid modifications transpire while tethered to CymA.

Bioinformatic analysis of CymV suggested that this cytochrome P450 is involved in olefin epoxidation. CymV shows sequence similarity to SalD (42% identity; ABG02269) and LtxB (31% identity; AAT12284), which are putative epoxidases involved in the biosynthesis of salinomycin[32] and lyngbyatoxin,[9]

(31) Chen, H. W.; Thomas, M. G.; O'Conner, S. E.; Hubbard, B. K.; Burkart, M. D.; Walsh, C. T. *Biochemistry* **2001**, *40*, 11651−11659.

respectively. To probe the function of CymV as the oxygenase responsible for the epoxidation of the isoprene olefin of **3** to **1**, we similarly inactivated the encoding gene *cymV* by PCR-based mutagenesis. HPLC-MS analysis of the resulting knockout mutant extract showed the loss of **1** with concomitant formation of **3**, which was produced at trace levels in the wild type (Figure 5). The structure of the mutant-derived **3** was verified by NMR comparison with published spectra.[1] Since **3** is produced in the mutant at levels comparable to the production of **1** in the wild-type strain, the CymV-catalyzed installation of the epoxide group probably occurs post-NRPS assembly. The wild-type and mutant strains both synthesize the olefinic **4**−**6** as anticipated.

The remaining two oxygenases CymS and CymW are homologous to cytochrome P450s and α-ketoglutarate-dependent iron dioxygenases, respectively, which hydroxylate aliphatic residues. CymS is homologous to CYP107Z4 (51% identity, AAT45275) from *Streptomyces lydicus*, which converts avermectin to 4′-oxo-avermectin at an aliphatic position.[33] CymW has 41% identity to a characterized L-proline 4-hydroxylase[34] (accession BAA20094) from a member of the *Dactylosporangium* genus. These enzymes are likely involved in the two remaining oxidation reactions involving the β-hydroxylation of tryptophan and stereospecific δ-hydroxylation of leucine (Figure 2).

Discussion

In this work we employed a multidisciplinary approach involving natural product discovery, stable isotope incorporation studies, genomics, and in vivo mutagenesis to unravel the biosynthesis of the related cyclomarazine dipeptides and the cyclomarin heptapeptides from two unrelated marine actinomycetes. These complementary experiments revealed an intriguing biosynthetic mechanism to nonribosomal peptides involving new pathways to nonproteinogenic amino acid residues as well as the in vivo formation of differently sized cyclic peptide products from the same NRPS assembly line.

The discovery of new cyclomarin analogs from *S. arenicola* CNS-205 and incorporation of stable isotopes into **1** from *S.* sp. CNB-982 were the starting point for developing biosynthetic proposals that were further evaluated genetically upon completion of the genome sequence of *S. arenicola* CNS-205. Bioinformatic and in vivo mutagenesis experiments correlated the *cym* locus to cyclomarin and cyclomarazine biosynthesis, thereby providing the molecular basis for the assembly of their amino acid components and their cyclic structures.

The molecular architecture of the NRPS megasynthetase CymA perfectly correlates with the synthesis of the cyclomarin heptapeptides starting with the tryptophan residue and ending with ADH. The majority of the cyclomarin building blocks are modified proteinogenic amino acids that have undergone methylation, oxidation, and prenylation reactions catalyzed by tailoring enzymes encoded by genes flanking the massive 23 kb *cymA* gene. Tryptophan, which is the priming amino acid residue, undergoes a series of modification events involving reverse *N*-prenylation by CymD, β-hydroxylation presumably by CymS or CymW, and epoxidation by CymV to yield *N*-(1,1-dimethyl-2,3-epoxypropyl)-β-hydroxytryptophan as in **1**. The timing of these biosynthetic events relative to peptide synthesis was probed in vivo, suggesting that prenylation and β-hydroxylation occur either on the free amino acid or while it is covalently tethered via a thioester linkage to CymA, whereas epoxidation occurs post NRPS assembly. Precedence for the prenylation reaction, however, suggested that the reaction would rather occur post NRPS assembly based on the biosynthesis of other prenylated NRPS products such as terrequinone A[35] and lyngbyatoxin.[9] Gene inactivation of the PTase *cymD* eliminated the production of the natural cyclomarins and cyclomarazines in *S. arenicola*, thereby establishing the biosynthetic linkage of these differently sized peptides. Furthermore, production of desprenylcyclomarin C (**7**) at a substantially lower concentration in the mutant relative to the corresponding cyclomarin and cyclomarazines in the wild-type organism suggested that prenylation is an integral biosynthetic reaction and that it must occur prior to the stage of the diketide intermediate common to the cyclomarazines and cyclomarins. In contrast, inactivation of the cytochrome P450 epoxidase *cymV* resulted in a mutant that efficiently shifted cyclomarin production from the epoxide **1** to the olefin **3**, which indicates that the liberated **3** is the substrate of CymV. This mutant also produced cyclomarazine **5**, which suggests that the diketide is not an in vivo substrate of CymV.

Intriguingly, the co-occurrence of the two cyclic peptide series derived from the same NRPS system is to the best of our knowledge unprecedented in nature. The key structural difference between the common amino acid residues in these peptides is the oxidation state of the β-position of the tryptophan moiety, which is oxidized in the heptapeptides and unoxidized in the diketide DKPs. DKPs have been reported to derive from NRPSs as well as NRPS-independent systems.[36] However, the majority of those deriving from NRPSs harbor a proline residue at the second amino acid position and are proposed to form spontaneously resulting from conformational constraints induced by the proline residue.[37,38] Few examples exist contrary to the necessity of proline for DKP formation by NRPSs, and in those cases,[39,40] the associated NRPS is programmed to produce a diketide. Whether formation of the cyclomarazine DKPs is facilitated by the type II TE CymQ serving an editing function to hydrolyze the incompletely processed dipeptide or its cleavage is nonenzymatic due to ineffective further processing by module 3 of the NRPS CymA is not yet clear (Figure 3) and the subject of further in vivo and in vitro studies. By regulating expression or activity of the tryptophan β-hydroxylase, production may be driven toward the selective synthesis of the cyclomarins or cyclomarazines.

A further unusual feature identified in this study involves the unexpected formation of the nonproteinogenic amino acid residue ADH. While our initial isotope labeling experiments led us astray in the formulation of a biosynthetic model,

(32) Izumikawa, M.; Murata, M.; Tachibana, K.; Ebizuka, Y.; Fujii, I. *Bioorg. Med. Chem.* **2003**, *11*, 3401−3405.
(33) Jungmann, V.; Molnar, I.; Hammer, P. E.; Hill, S.; Zirkle, R.; Buckel, T. G.; Buckel, D.; Ligon, J. M.; Pachlatko, J. P. *Appl. Environ. Microbiol.* **2005**, *71*, 6968−6976.
(34) Shibasaki, T.; Mori, H.; Chiba, S.; Ozaki, A. *Appl. Environ. Microbiol.* **1999**, *65*, 4028−4031.
(35) Balibar, C. J.; Howard-Jones, A. R.; Walsh, C. T. *Nat. Chem. Biol.* **2007**, *3*, 584−592.
(36) Lautru, S.; Gondry, M.; Genet, R.; Pernodet, J. L. *Chem. Biol.* **2002**, *9*, 1355−1364.
(37) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. *J. Biol. Chem.* **1998**, *273*, 22773−22781.
(38) Stachelhaus, T.; Walsh, C. T. *Biochemistry* **2000**, *39*, 5775−5787.
(39) Healy, F. G.; Wach, M.; Krasnoff, S. B.; Gibson, D. M.; Loria, R. *Mol. Microbiol.* **2000**, *38*, 794−804.
(40) Balibar, C. J.; Walsh, C. T. *Biochemistry* **2006**, *45*, 15029−15038.

bioinformatic analysis of the *cym* locus suggested an alternative pathway related to the degradation of 3-HPP, yet operating in the reverse direction (Figure 4). The acquisition and evolution of the ADH biosynthetic pathway represents an intriguing example of pathway creation, which is being further explored biochemically.

Experimental Section

General Experimental Procedures. Low-resolution LC/MS data was acquired using a Hewlett-Packard series 1100 LC/MS system utilizing electrospray ionization in positive mode with a linear gradient of 10−90% aqueous MeCN at a flow rate of 0.7 mL/min over 24 min on a reversed-phase C_{18} column (Phenomenex Luna, 4.6 mm × 100 mm, 5 μm). High-resolution mass spectra were collected by HR-ESI-TOFMS at the Scripps Research Institute, La Jolla, CA, and ESI-HR-FTMS at the Chemistry & Biochemistry Mass Spectrometry Facility, University of California San Diego. NMR spectral data were obtained on Varian Inova 300 and 500 MHz spectrometers and Bruker AM400 and DRX600 spectrometers. Optical rotations were measured using a Rudolph Research Autopol III polarimeter with a 10 cm cell. UV spectra were recorded in a Varian Cary UV−visible spectrophotometer with a path length of 1 cm. CD spectra were collected in an AVIV model 215 CD spectrometer with a 0.5 cm long cell. IR spectra were acquired in a Nicolet Magna 550 FT-IR series II spectrometer.

Bacterial Strains, Plasmids, Culture Conditions, and DNA Manipulations. The marine actinomycete strain CNS-205 was isolated from a sediment sample collected at a depth of 20 m off Palau in 2004. The strain was identified as *S. arenicola* based on 16S rDNA analysis (100% identity; NC_009953).[17] *Streptomyces* sp. CNB-982 was previously described.[1] Both bacteria were grown in A1 media (10 g of starch, 4 g of yeast extract, 2 g of peptone, and 10 mL of 1 M Tris (pH 8.0) per liter of seawater) at 28 °C and 200 rpm unless otherwise noted.

Escherichia coli DH5α (Invitrogen) was used for cloning experiments as described,[41] and *E. coli* S17-1[42] and ET12567/pUZ8002[43] were used for conjugation experiments. Previously adapted for use in *Salinispora*,[30] REDIRECT (Plant Bioscience Limited, Norwich, U.K.) technology was utilized for PCR targeting.[44] The λ-Red function was provided by pKD20 (bla).[45] Fosmid BPPW9569 containing a fragment of the *cym* gene cluster from *cymS* and a transposase directly downstream of *cymH* in the pCC1FOS vector (Epicentre) was obtained from the Joint Genome Institute, Walnut Creek, CA, and utilized for both *cymD* and *cymV* deletions. Apramycin (100 μg/mL for *S. arenicola*; 50 μg/mL for *E. coli*), chloramphenicol (2.5 μg/mL for *S. arenicola*; 12.5−25.0 μg/mL for *E. coli*), carbenicillin (100 μg/mL), and nalidixic acid (100 μg/mL) were used for selection of recombinants. DNA purification and manipulation was performed according to standard procedures.[41,46]

Cultivation and Isolation of 1 and 3−6. *S. arenicola* CNS-205 was cultured at 27 °C with shaking at 250 rpm in 20 2.8 L Fernbach flasks each containing 1 L of the medium A1BFe+C (10 g of starch, 4 g of yeast extract, 2 g of peptone, 1 g of $CaCO_2$, 40 mg of $Fe_2(SO_4)_3 \cdot 4H_2O$, 100 mg of KBr, and 1 L of seawater). After 13 days, organic chemical constituents from 20 L culture were collected by solid-phase extraction using Amberlite XAD-7 resin. The crude extract was prepared by washing the resin with acetone and evaporating the solvent in vacuo.

(41) Sambrook, J.; Russell, D. W. *Molecular Cloning, a Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.
(42) Simon, R.; Priefer, U.; Pühler, A. *Bio/Technology* **1983**, *1*, 784−791.
(43) Paget, M. S. B.; Chamberlin, L.; Atrih, A.; Foster, S. J.; Buttner, M. J. *J. Bacteriol.* **1999**, *181*, 204−211.
(44) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1541−1546.
(45) Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6640−6645.
(46) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, 2000.

The crude extract (14 g) was partitioned with EtOAc and H_2O. The EtOAc layer was dried under vacuum to yield 2 g of crude material, which was fractionated through a C_{18} vacuum column (4.5 cm diameter × 10 cm height) by eluting with a step gradient from 20%, 40%, 60%, 80%, to 100% MeOH in H_2O. The 60% MeOH/H_2O fraction (163 mg) was subjected to reversed-phase HPLC with a gradient of aqueous MeCN (0−7 min, 20%; 7−30 min, 20−35%; 25−45 min, 35% MeCN) in a preparative LC (Waters Prep LC 4000 system, Waters preparative column C_{18} 25 mm × 200 mm, 10 mL/min, UV detection at 210 nm). Cyclomarazines A and B (**5** and **6**) eluted at 41 and 37 min, respectively, yielding 0.5 mg of **6**. Cyclomarazine A was further purified by reversed-phase HPLC (Alltech semipreparative column C_{18}, 10 mm × 250 mm, 2 mL/min, refractive index detection) with the isocratic solvent of 35% MeCN in H_2O to yield 10 mg of **5** eluting at 29 min. The 80% MeOH/H_2O fraction was further fractionated with a preparative LC (Waters Prep LC 4000 system, Waters preparative column C_{18} 25 mm × 200 mm, 10 mL/min, UV detection at 210 nm) using a gradient solvent of MeCN/H_2O (0−15 min, 50%; 15−25 min, 50−73%; 25−40 min, 73% MeCN). Cyclomarin D, which eluted at 21 min, was further purified with reversed-phase HPLC (Alltech semipreparative column C_{18}, 10 mm × 250 mm, 2 mL/min, refractive index detection) with 75% MeCN/H_2O, yielding 22 mg of **4** at 21 min. In a similar fashion from a 1 L culture of *S. arenicola* CNS-205, 3.2 mg of **1** and 0.3 mg of **3** (eluting at 21 and 25 min, respectively) were isolated.

Cyclomarin D (4). Oil; $[\alpha]_D$ −30 (c = 0.80, $CHCl_3$); IR (neat) ν_{max} 3329, 2962, 2871, 1638, 1510 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 221 (4.5), 283 (3.7) nm; CD (MeCN) ($\Delta\epsilon$) 232 (+4.8), 210 (−25.4); ^{1}H NMR (500 MHz, $CDCl_3$) δ (assignment, multiplicity, coupling constants in Hertz, *COSY*s, **HMBC**s): 0.65 (H24, d, J = 6.5, *H19*, **C18**, **C19**, **C20**), 0.82 (H23, s, **C21**), 0.89 (H54, d, J = 6.5, *H52*, **C52**), 0.95 (H30, d, J = 6.5, *H28*, **C27**, **C28**, **C29**), 0.97 (H35, d, J = 6.5, *H33*, **C33**), 1.00 (H34, d, J = 6.5, *H33*, **C33**), 1.02 (H29, d, J = 6.5, *H28*, **C27**, **C28**, **C30**), 1.35 (H48, d, J = 7.5, *H47*, **C46**, **C47**), 1.40 (H22, s, **C20**, **C21**, **C23**), 1.66 (H28, m, *H27, H29, H30*), 1.76 (H16, s, **C12**, **C13**, **C14**, **C15**), 1.77 (H15, s, **C12**, **C13**, **C14**, **C16**), 1.83 (H52, m, *H51, H53, H54*), 1.91 (H51b, m, *H50, H51a, H52*), 1.91 (H27b, m, *H26, H27a, H28*), 1.97 (H51a, m, *H50, H51b, H52*, **C50**, **C52**, **C53**, **C54**), 2.07 (H27a, m, *H26, H27b, H28*, **C50**, **C52**, **C53**, **C54**), 2.21 (H33, m, *H32, H34, H35*, **C32**, **C34**, **C35**), 3.20 (H19, m, *H18, H20, H24*), 3.27 (26-NMe, s, **C20**, **C31**), 3.30 (H34, s, **C38**), 3.45 (H53b, m, *H52, H53b*, **C51**, **C52**, **C54**), 3.54 (H53a, dd, J = 10.0, 9.5, *H53b, H52*, **C51**, **C52**, **C54**), 3.55 (H26, m, *H27a, H27b*, **C25**, **C27**, **C28**), 3.62 (H18, br m, *18-NH, H19*) 3.83 (H47, m, *47-NH, H48*), 4.34 (H37, br d, J = 6.5, *37-NH, H38*, **C36**, **C38**), 4.62 (H32, m, *32-NH, H33*, **C31**), 4.64 (H20, d, J = 9.0, *H19*), 4.70 (H2, br m, *2-NH, H3*), 4.76 (H50, br m, *50-NH, 51*, **C49**), 4.94 (H38, br m, *H37*, **C37**, **C39**, **C40**, **C44**), 5.19 (H14b, d, J = 17.5, *H13*, **C12**, **C13**, **C15**, **C16**), 5.21 (H14a, d, J = 10.5, *H13*, **C12**, **C13**, **C15**, **C16**), 5.63 (H3, d, J = 2.5, *H2*), 6.12 (H13, dd, J = 17.0, 10.5, *H14b, H14a*, **C12**, **C15**, **C16**), 6.75 (18-NH, br m, *H18*), 7.05 (H8, dd, J = 7.5, 7.5, *H7, H9*, **C6**, **C10**), 7.09 (H9, dd, J = 8.0, 7.5, *H9, H10*, **C7**, **C11**), 7.21 (32-NH, m, *H32*), 7.31 (50-NH, m, *H50*), 7.34 (H40, m, *H41*, **C42**, **C44**), 7.34 (H44, m, *H43*, **C40**, **C42**), 7.35 (H41, m, *H40, H42*, **C39**, **C43**), 7.35 (H42, m, *H41, H43*, **C40**, **C44**), 7.35 (H43, m, *H42, H44*, **C39**, **C41**), 7.46 (H10, d, J = 8.0, *H9*, **C6**, **C8**), 7.53 (37-NH, br m, *H37*), 7.58 (H4, s, **C2**, **C3**, **C5**, **C6**, **C11**, **C12**), 7.67 (H7, d, J = 7.5, *H8*, **C9**, **C11**), 7.79 (47-NH, br m, *H47*), 8.15 (2-NH, br m, *H2*); ^{13}C NMR (75 MHz, $CDCl_3$) δ 16.9 (C23), 17.3 (C24), 17.3 (C48), 17.9 (C54), 17.9 (C34), 19.6 (C35), 21.9 (C29), 23.1 (C30), 25.5 (C22), 25.5 (C28), 27.6 (C15), 27.9 (C16), 30.6 (C33), 32.4 (C52), 32.4 (C19), 35.5 (C51), 38.3 (C27), 40.3 (26-NMe), 51.0 (C50), 52.5 (C47), 55.5 (C32), 57.7 (C45), 58.9 (C2), 59.1 (C12), 60.6 (C37), 62.5 (C18), 65.4 (C26), 67.4 (C53), 67.7 (C3), 81.3 (C38), 113.1 (C5), 113.5 (C14), 113.8 (C10), 118.9 (C7), 119.2 (C8), 121.0 (C9), 123.5 (C4), 125.1 (C20), 126.3 (C40), 126.3 (C44), 127.0 (C6), 128.2 (C42), 128.5 (C41), 128.5 (C43), 133.8 (C21), 135.9 (C11), 137.5 (C39), 144.4 (C13), 169.8

(C36), 171.2 (C1), 171.2 (C25), 172.7 (C31), 173.0 (C46), 173.9 (C17), 173.9 (C49); HR-ESI-TOFMS $[M + Na]^+$ *m/z* 1035.5909 ($C_{55}H_{80}N_8O_{10}$, calcd $[M + Na]^+$ 1035.5889).

Cyclomarazine A (5). Oil; $[\alpha]_D$ −9 (*c* = 0.19, MeOH); IR (neat) ν_{max} 3431, 2960, 1647, 1528 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 225 (3.7), 285 (3.0) nm; CD (MeCN) ($\Delta\epsilon$) 234 (+1.1), 218 (−2.6); ^1H NMR (300 MHz, DMSO-d_6) δ 0.59 (H19b, m, *H18*, *H19a*, *H20*, **C17**, **C18**, **C20**, **C21**, **C22**), 0.60 (H22, d, *J* = 6.5, *H20*, **C19**, **C20**, **C21**), 0.68 (H19a, ddd, *J* = 12.0, 9.0, 3.5, *H19b*, *H18*, *H20*, **C17**, **C18**, **C20**, **C21**, **C22**), 1.40 (H20, m, *H19a*, *H19b*, *H21a*, *H21b*, **C19**, **C21**, **C22**), 1.67 (H15, s, **C4**, **C12**, **C13**, **C14**, **C16**), 1.67 (H16, s, **C4**, **C12**, **C13**, **C14**, **C15**), 2.71 (18-NMe, s, **C1**, **C18**), 2.77 (H21b, br m, *H20*, *H21a*, *21-OH*, **C22**), 2.84 (H21a, br m, *H20*, *H21b*, *H21-OH*, **C20**, **C22**), 3.05 (H3b, dd, *J* = 14.5, 5.5, *H3a*, *H2*, **C1**, **C2**, **C4**, **C5**, **C6**), 3.15 (H3a, dd, *J* = 14.5, 5.5, *H3b*, *H2*, **C1**, **C2**, **C4**, **C5**, **C6**), 3.61 (H18, dd, *J* = 9.0, 3.5, *H19a*, *H19b*, **C1**, **C17**, **C19**, **C20**), 4.13 (H2, br m, *2-NH*, *H3a*, *H3b*, **C1**, **C5**), 4.31 (21-OH, br m, *H21a*, *H21b*), 5.18 (H14b, d, *J* = 17.0, *H13*, **C12**, **C13**), 5.19 (H14a, d, *J* = 11.0, *H13*, **C12**, **C13**), 6.09 (H13, dd, *J* = 17.0, 11.0, *H14b*, *H14a*, **C12**, **C15**, **C16**), 6.96 (H8, dd, *J* = 8.0, 8.0, *H7*, *H9*, **C6**, **C10**), 7.01 (H9, dd, *J* = 8.0, 8.0, *H8*, *H10*, **C7**, **C11**), 7.15 (H4, s, **C3**, **C5**, **C6**, **C11**, **C12**), 7.41 (H10, d, *J* = 8.0, *H9*, **C6**, **C8**), 7.51 (H7, d, *J* = 8.0, *H8*, **C5**, **C6**, **C9**, **C11**), 8.23 (2-NH, d, *J* = 2.5, *H2*); ^{13}C NMR (125 MHz, DMSO-d_6) δ 17.0 (C22), 28.0 (C15), 28.0 (C16), 30.8 (C3), 32.5 (18-NMe), 33.1 (C20), 36.9 (C19), 56.5 (C2), 58.0 (C12), 59.3 (C18), 66.5 (C21), 108.5 (C5), 114.0 (C10), 114.0 (C14), 119.0 (C8), 119.7 (C7), 121.7 (C9), 125.7 (C4), 130.1 (C6), 135.5 (C11), 144.6 (C13), 166.3 (C1), 167.9 (C17); HR-ESI-TOFMS $[M + H]^+$ *m/z* 398.2435 ($C_{23}H_{31}N_3O_3$, calcd $[M + H]^+$ 398.2438).

Cyclomarazine B (6). Oil; $[\alpha]_D$ −16 (*c* = 0.03, MeOH); IR (neat) ν_{max} 3439, 2960, 1655, 1519 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 224 (3.6), 284 (3.0) nm; CD (MeCN) ($\Delta\epsilon$) 232 (+1.4), 218 (−2.3); ^1H NMR (500 MHz, DMSO-d_6) δ 0.00 (H19b, ddd, *J* = 13.5, 8.5, 5.0, *H19a*, *H18*, *H20*, **C17**, **C18**, **C20**, **C21**, **C22**), 0.33 (H22, d, *J* = 6.5, *H20*, **C19**, **C20**, **C21**), 1.09 (H19a, ddd, *J* = 13.5, 9.0, 2.5, *H19a*, *H18*, *H20*, **C17**, **C18**, **C20**, **C21**, **C22**), 1.25 (H20, m, *H19a*, *H19b*, *H21*, **C19**, **C21**, **C22**), 1.66 (H16, s, **C12**, **C13**, **C14**, **C15**), 1.67 (H15, s, **C12**, **C13**, **C14**, **C15**), 2.90 (H3b, dd, *J* = 14.5, 5.5, *H3a*, *H2*, **C1**, **C2**, **C4**, **C5**, **C6**), 2.93 (H21, br m, *H20*, *H21-OH*, **C19**, **C20**, **C22**), 3.24 (H3a, dd, *J* = 14.5, 5.5, *H3b*, *H2*, **C1**, **C2**, **C4**, **C5**, **C6**), 3.57 (H18, ddd, *J* = 9.0, 3.5, 1.0, *H19a*, *H19b*, *18-NH*, **C17**), 4.09 (H2, br m, *2-NH*, *H3a*, *H3b*, **C1**, **C5**), 4.30 (21-OH, br m, *H21*), 5.16 (H14b, d, *J* = 17.5, *H13*, **C12**, **C13**), 5.19 (H14a, d, *J* = 10.5, *H13*, **C12**, **C13**), 6.08 (H13, dd, *J* = 17.0, 10.5, *H14b*, *H14a*, **C12**, **C15**, **C16**), 6.93 (H8, dd, *J* = 7.5, 7.5, *H7*, *H9*, **C6**, **C10**), 6.98 (H9, dd, *J* = 8.0, 7.5, *H8*, *H10*, **C7**, **C11**), 7.18 (H4, s, **C3**, **C5**, **C6**, **C11**, **C12**), 7.39 (H10, d, *J* = 8.0, *H9*, **C6**, **C8**), 7.54 (H7, d, *J* = 7.5, *H8*, **C5**, **C6**, **C9**, **C11**), 7.82 (2-NH, d, *J* = 1.0, *H2*, **C2**, **C17**), 8.07 (18-NH, d, *J* = 1.0, *H18*, **C1**, **C18**); ^{13}C NMR (75 MHz, DMSO-d_6) δ 17.2 (C22), 27.4, (C16), 27.4 (C15), 29.0 (C3), 30.8 (C20), 39.7 (C19), 52.4 (C18), 55.2 (C2), 58.5 (C12), 65.2 (C21), 107.3 (C5), 113.1 (C10), 113.3 (C14a), 118.4 (C8), 119.2 (C7), 120.3 (C9), 125.7 (C4), 129.5 (C6), 134.7 (C11), 144.0 (C13), 167.0 (C1), 167.7 (C17); HR-ESI-TOFMS $[M + H]^+$ *m/z* 384.2281 ($C_{22}H_{29}N_3O_3$, calcd $[M + H]^+$ 384.2282).

Flash Acid Hydrolysis and Advanced Marfey Analysis of 5 and 6. One milligram of **5** was hydrolyzed in 0.5 mL of 6 N HCl at 110 °C for 1 h, and HCl was then evaporated under vacuum. The dry material was resuspended in 0.5 mL of H_2O and dried three times by lyophilization to remove residual acid. The hydrolysate was loaded onto a C_{18} Sepak column (0.5 cm diameter, 1 cm height) and eluted with 10% MeCN in H_2O. The purified hydrolysate was divided into two portions and dissolved in 1 N $NaHCO_3$ (100 μL) followed by adding 50 μL of 3 mg/mL l-1-fluoro-2,4,-dinitrophenyl-5-leucine amide (L-FDLA) and D-FDLA, respectively. The reaction mixture was incubated at 80 °C for 3 min and then quenched by neutralization with 50 μL of 2 N HCl. A 300 μL amount of 50% aqueous MeCN was added to the solution to dissolve the reaction mixture. The products were analyzed by LC/MS with a gradient solvent system from 30% to 80% MeCN (0.1% TFA) over 50 min *N*-(1,1-dimethyl-allyl)-tryptophan was eluted at retention times of 13.6 and 14.1 min with L-FDLA and D-FDLA derivatizations, respectively. *N*-Methyl-δ-hydroxyleucine was detected at 13.1 and 15.6 min with L-FDLA and D-FDLA, respectively. These results allowed the absolute configurations of the α-carbons of both amino acid residues to be assigned as 2*S* and 18*S*. For **6**, the same procedure was repeated. L-FDLA and D-FDLA derivatives of δ-hydroxyleucine were eluted at 12.7 and 15.2 min, respectively.

Flash acid hydrolysis and advanced Marfey analysis were also applied for cyclomarin D and A (**4** and **1**). The retention times for 2-amino-3,5-dimethylhex-4-enoic acid, *N*-methylleucine, valine, β-methoxyphenylalanine, and alanine from **4** were 22.4, 23.9, 18.1, 21.2, and 14.7 min with L-FDLA derivatization and 26.3, 27.3, 24.7, 27.8, and 18.2 min with D-FDLA derivatization, respectively. The retention times of the derivatives of 2-amino-3,5-dimethylhex-4-enoic acid and β-methoxyphenylalanine from **1** were consistent with those from analysis of **4**.

Labeling Experiments. Initial feeding experiments were carried out in cultures of *Streptomyces* sp. CNB-982. Production cultures in A1 media were incubated in Fernbach flasks for 6−7 days at 230 rpm and 23 °C. A mixture of 1 g of L-[U-$^{13}C_6$]glucose and 1 g of D-glucose was pulse fed at 48, 62, 73, 88, and 96 h to a 1 L production culture. Pulse feeding [methyl-^{13}C]methionine (52.5 mg) at 48, 60, 72, 85, and 94 h and D,L-[3′-^{13}C]tryptophan (90.1 mg) at 40, 52, 64, 76, and 88 h to two separate 1 L production cultures was also carried out. An additional feeding experiment was carried out in cultures of *S. arenicola* CNS-205. Production cultures consisting of A1+BFe media (1 L) were incubated in Fernbach flasks for 10 days at 230 rpm and 28 °C. Sodium [1-^{13}C]isobutyrate[47] (25 mg) was fed at 60 h.

Isolation of Cyclomarin A from *Streptomyces* sp. CNB-982. Cells from a 1 L fermentation were removed by centrifugation at 10 000*g* for 10 min. The supernatant was extracted with EtOAc (2 × 800 mL), and the upper layer was dried over anhydrous Na_2SO_4 and evaporated to dryness to give 35−55 mg of residue. The extract was suspended in MeOH/H_2O (8:2) and passed through a 500 mg ODS Bond-Elut column, eluting with 3 column volumes of the same solvent. Purification of the residue by HPLC [10 × 250 mm YMC-AQ ODS column, MeOH/H_2O (70:30), 5 mL/min] yielded 3−10 mg of **1**.

DNA Sequence Analysis of the Cyclomarin Gene Cluster (cym). The putative cyclomarin NRPS *cymA* was identified from a draft genome assembly of *S. arenicola* CNS-205 (GenBank accession CP00850) provided by the Joint Genome Institute (JGI) through the Community Sequencing Program. NRPS domain structure for *cymA* was determined as described,[48] and adenylation domain specificity was predicted using NRPSPredictor.[49] Sequence analysis and putative functions for enzymes encoded by genes surrounding *cymA* were predicted using the suite of tools available at NBCI (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). Open reading frames were further analyzed by FramePlot.[50]

Conjugation Protocol for *S. arenicola*. *S. arenicola* CNS-205 was cultivated for 5 days in 50 mL of A1 in a 250 mL flask containing a stainless steel spring. Mycelium was harvested from 25 mL and resuspended in 2 mL of A1. A 0.5 mL amount of *E. coli* S17-1 suspension[46] was mixed with 0.5 mL of the *S. arenicola* suspension and then plated on A1 agar plates. After 1 day incubation at 34 °C, the plates were overlaid with 2.5 mg each of apramycin and nalidixic acid. Exconjugates were visible 10 days after plate inoculation.

(47) Moore, B. S.; Seng, D. *Tetrahedron Lett.* **1998**, *39*, 3915−3918.
(48) Udwary, D. W.; Zeigler, L.; Asolkar, R. N.; Singan, V.; Lapidus, A.; Fenical, W.; Jensen, P. R.; Moore, B. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10376−10381.
(49) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. *Nucleic Acids Res.* **2005**, *33*, 5799−5808.
(50) Ishikawa, J.; Hotta, K. *FEMS Microbiol. Lett.* **1999**, *174*, 251−253.

Inactivation of *cymD* and *cymV*. *cymD* was inactivated utilizing the modified PCR targeting system previously described for *S. tropica*.[30] In fosmid BPPW9569, *cymD* was replaced by the apramycin resistance (*acc(3)IV*) cassette from pIJ773 utilizing λ-Red-mediated recombination. The *cymD* replacement cassette was generated by PCR using primers cymDkoF1 (5′-cttccgacgacacgacgacggggcggcgtgagcggtatgATTCCGGGGATCCGTCGACC-3′) and cymDkoR1 (5′-gatcgcgtgacgtgcggcgaccggcgaggtgcccgatcaTGTAGGCTGGAGCTGCTTC-3′). Lowercase letters represent 39 nt homologous extensions immediately upstream and downstream of *cymD*, respectively, including the putative start and stop codons. This cassette was introduced into *E. coli* BW25113/pKD20[45] containing the chloramphenicol-resistant fosmid BPPW9569. Gene replacement was verified using PCR primers cymDckF1 (5′-agcgtgagatcggcctaggg-3′) and cymDckR1 (5′-TTGTGGTAGCTCGCGAGCTG-3′). The mutated fosmid was introduced into *S. arenicola* CNS-205 by conjugation from *E. coli* S17-1 and confirmed by PCR analysis.

An identical methodology was utilized to inactivate *cymV*. The *cymV* replacement cassette was generated using primers cymVkoF1 (5′-cgcgcctcgacgtggattctgacaaggagaaccgggttgATTCCGGGGATCCGTCGACC-3′) and cymVkoR1 (5′-gacgtcgtgcatctccggtgacaggtagaccttcatccgTGTAGGCTGGAGCTGCTTC-3′). To verify replacement of *cymV* with *acc-(3)IV* in fosmid BPPW9569 and *S. arenicola* CNS-205, the primers cymVckF1 (5′-tggttcccggccgtgatcgg-3′) and cymVckR1 (5′-CGGCGAACAGGACGCACGTC-3′) were utilized.

Analysis of the *S. arenicola cymD*$^-$ and *cymV*$^-$ Mutants. A 100 μL amount of 4 day cultures of the *S. arenicola cymD*$^-$ and *cymV*$^-$ knockout mutants was used to inoculate 100 mL of A1+BFe in a 500 mL flask containing a stainless steel spring. Apramycin (100 μg/mL) was added to cultures of *cymD*$^-$ and *cymV*$^-$ at the time of inoculation. Cultivation was carried out for 10 days at 28 °C and 200 rpm, and the cultures were extracted and analyzed by LC/MS as described above. Larger scale fermentations of the *cymV*$^-$ (1 L) and *cymD*$^-$ (10 L) mutants were cultured and extracted, and the products were isolated as described above. The *S. arenicola cymV* knockout mutant yielded 2.1 mg of **3**, whereas the *cymD*$^-$ mutant gave 0.5 mg of a crude fraction containing **7** (HR-FTMS $[M + Na]^+$ *m*/*z* 981.5408, calcd 981.5420).

Acknowledgment. We thank Dr. A. S. Eustáquio (UCSD) for valuable discussions regarding PCR-targeted mutagenesis and Dr. J. A. Kalaitzis (UCSD) for assistance with NMR experiments. This research was supported in part by NOAA through the Oceans and Human Health Initiative (NA05NOS-4781249 to B.S.M. and P.R.J.), the National Institutes of Health (CA44848, W.F.), and the California and Oregon Sea Grant Programs (85-NMP-N to B.S.M. and P.R.J.; NA36RG0451 to S.J.G.). Genome sequencing was provided through a grant from the Community Sequencing Program of JGI to P.R.J. and B.S.M.

Supporting Information Available: Complete ref 10, 1D and 2D NMR spectra (^{1}H, ^{13}C, gCOSY, TOCSY, ROESY, gHSQC, gHMBC) for **4**−**6**, ^{13}C and INADEQUATE NMR spectra for ^{13}C-labeled **1**, NMR tables, and a CD spectrum for **5** and **6** (41 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

JA711188X