Cloning and Biochemical Characterization of the Hectochlorin Biosynthetic Gene Cluster from the Marine Cyanobacterium *Lyngbya majuscula*

Aishwarya V. Ramaswamy,[†] Carla M. Sorrels,[‡,§] and William H. Gerwick*[,‡,§,⊥]

Department of Microbiology, Oregon State University, Corvallis, Oregon 97331, College of Pharmacy, Oregon State University, Corvallis, Oregon 97331, Center for Marine Biomedicine and Biotechnology, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92037, and Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92037

Received August 14, 2007

Cyanobacteria, or blue-green algae, are a rich source of novel bioactive secondary metabolites that have potential applications as antimicrobial or anticancer agents or useful probes in cell biology studies. A Jamaican collection of the cyanobacterium *Lyngbya majuscula* has yielded several unique compounds including hectochlorin (**1**) and the jamaicamides A−C (**5**–**7**). Hectochlorin has remarkable antifungal and cytotoxic properties. In this study, we have isolated the hectochlorin biosynthetic gene cluster (*hct*) from *L. majuscula* to obtain details regarding its biosynthesis at the molecular genetic level. The genetic architecture and domain organization appear to be colinear with respect to its biosynthesis and consists of eight open reading frames (ORFs) spanning 38 kb. An unusual feature of the cluster is the presence of ketoreductase (KR) domains in two peptide synthetase modules, which are predicted to be involved in the formation of the two 2,3-dihydroxyisovaleric acid (DHIV) units. This biosynthetic motif has only recently been described in cereulide, valinomycin, and cryptophycin biosynthesis, and hence, this is only the second such report of an embedded ketoreductase in a cyanobacterial secondary metabolite gene cluster. Also present at the downstream end of the cluster are two cytochrome P450 monooxygenases, which are likely involved in the formation of the DHIV units. A putative halogenase, at the beginning of the gene cluster, is predicted to form 5,5-dichlorohexanoic acid.

The marine environment continues to offer a vast array of biologically active natural products and constitutes a major source of molecules for drug discovery and development.[1] Among these organisms, marine cyanobacteria have emerged as one of the richest producers of new natural product chemotypes with potent biological properties.[2–4] A predominant metabolic theme in cyanobacterial metabolites is the rich integration of polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS).[3] PKSs and NRPSs represent a large group of multifunctional enzymes with a modular organization that assemble complex molecules using acetate or amino acids as building blocks, respectively. Additionally, these secondary metabolites exhibit tremendous structural diversity due to the presence of numerous tailoring enzymes that catalyze oxidation, reduction, methylation, different types of halogenation, and epimerization.[5,6] A complete understanding of the biosynthesis of these compounds at the molecular genetic level will provide key information for harnessing these gene clusters to produce novel molecules through combinatorial genetic engineering.

Hectochlorin (**1**) was first isolated from a Jamaican isolate of the cyanobacterium *Lyngbya majuscula* (designated as *L. majuscula* strain JHB). It was subsequently reisolated from *L. majuscula* collected at Bocas del Toro in Panama. The planar structure of hectochlorin was elucidated by NMR and MS analyses, and its absolute configuration was deduced by X-ray crystallography.[7] Structurally, hectochlorin is related to dolabellin (**2**) isolated from the sea-hare *Dolabella auricularia* and lyngbyabellin A (**3**) and lyngbyabellin B (**4**) from *L. majuscula* (Figure 1).[8–11] Initial antimicrobial assays indicated that hectochlorin was a potent antifungal agent. Its fungicidal activity and potency against a number of plant pathogens initiated synthetic efforts, and a total synthesis of hectochlorin has been reported.[12] Further testing revealed that Ptk2 cells (derived from the rat kangaroo, *Potorous tridactylus*) treated with hectochlorin showed an increase in the number of binucleated cells as a result of arrest of cytokinesis, a result typical of compounds that cause hyperpolymerization of actin. Hectochlorin was also evaluated against the *in vitro* panel of 60 different cancer cell lines at the National Cancer Institute and showed potent activity toward cell lines in the colon, melanoma, ovarian, and renal subpanels. It yielded a flat dose–response curve against most cell lines, a characteristic of compounds that are antiproliferative but not directly cytotoxic.[7] The unique biological properties and presence of intriguing structural features in hectochlorin, such as the gem-dichloro group and two DHIV units, prompted us to clone and characterize its biosynthetic gene cluster. In this study, we report the cloning, sequencing, and partial biochemical characterization of the hectochlorin biosynthetic pathway from *L. majuscula* JHB.

Results and Discussion

Cloning and Identification of the Hectochlorin Biosynthetic Gene Cluster. The structure of hectochlorin (**1**) suggested that it was derived from a mixed PKS/NRPS pathway. C1–C8 were proposed to arise from intact acetate units assembled by a PKS with the pendent methyl group (C9) added from *S*-adenosylmethionine (SAM). Portions of the thiazole rings, C10–C12 and C18–C20, plus the heteroatoms were proposed to originate from cysteine.[13,14] DHIV, an intermediate of valine biosynthesis, was proposed to form C13–C17 and C21–C25 units. Therefore, our initial strategy to isolate the hectochlorin biosynthetic gene cluster involved the use of more general adenylation domain probes based on the A2 and A8 conserved motifs.[15] Unfortunately, this protocol proved unsuccessful in isolating the *hct* gene cluster and instead led to the identification of an approximately 20 kilobase (kb) cluster whose genetic architecture indicated its involvement in the biosynthesis of an unrelated metabolite (A. V. Ramaswamy, Ph.D. Thesis, Oregon State University, 2005).

Although chemical analysis of *L. majuscula* JHB has yielded only the jamaicamides A–C (**5**–**7**),[16] hectochlorin (**1**), and 23-deoxyhectochlorin (K. L. McPhail and W. H. Gerwick, unpublished data) as major metabolites, the presence of multiple PKS/NRPS

* To whom correspondence should be addressed. Tel: (858) 534-0578. Fax: (858) 534-0529. E-mail: wgerwick@ucsd.edu.
[†] Department of Microbiology.
[‡] College of Pharmacy.
[§] Scripps Institution of Oceanography.
[⊥] Skaggs School of Pharmacy and Pharmaceutical Sciences.

10.1021/np0704250 CCC: $37.00
Published on Web 11/15/2007

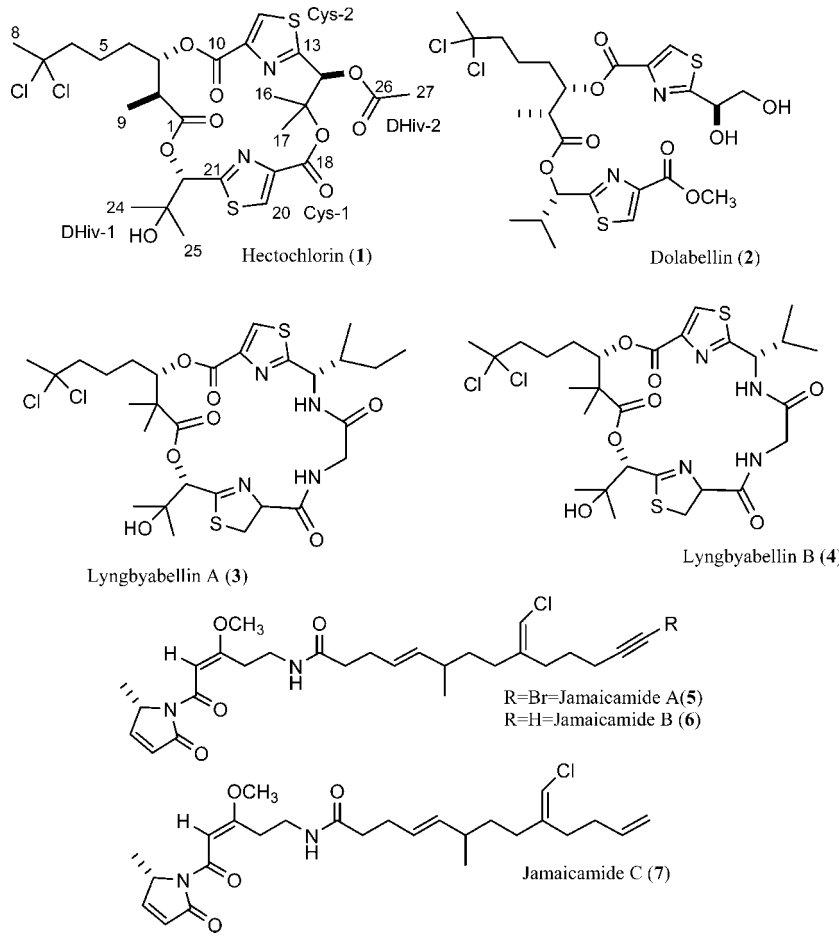

Figure 1. Chemical structure of hectochlorin (**1**) and related metabolites, dolabellin (**2**), lyngbyabellin A (**3**), and lyngbyabellin B (**4**). Jamaicamides A−C (**5**–**7**) are additional major metabolites isolated from *L. majuscula* JHB.

pathways was detected during library screening (D. J. Edwards and W. H. Gerwick, unpublished). Therefore, a more specific approach was used that focused on the modules involved in cysteine incorporation, and specifically on the heterocyclization domains within these modules. Degenerate primers spanning the heterocyclization domain and conserved region of the adenylation domain (between core motifs A4 and A5) were designed on the basis of alignments of previously characterized heterocyclization domains. These primers were used to amplify an approximately 1.7 kb fragment from *L. majuscula* JHB genomic DNA, which was cloned into pGEM-T Easy. Several clones were sequenced and yielded three distinct heterocyclization−adenylation domain products, thus indicating that there were three different gene clusters or three distinct modules involved in the activation and incorporation of cysteine in this *L. majuscula* genome.

An equal concentration of these three digoxigenin (DIG)-labeled probes was used to screen a library of approximately 2800 clones, which identified 12 different fosmids at the colony hybridization level. Restriction digest mapping with *Hin*dIII and *Eco*RI enzymes and Southern hybridization analyses using the heterocyclization−adenylation domain probe led to the identification of two overlapping fosmids, pHctI and pHctII, as candidates for the *hct* gene cluster. To determine if these fosmids contained PKS elements, a pair of degenerate primers based on the β-ketosynthase (KS) domain were used to amplify a 0.7 kb fragment from this domain.[17] Sequence analysis of the PCR products indicated that the two fosmids contained a single, identical KS domain. This result was consistent with the hypothesis that a hexanoic acid derivative was undergoing a single round of ketide extension in the initiation of the hectochlorin biosynthetic pathway, and this was a familiar biosynthetic logic from our prior work with jamaicamide A (**5**), in which hexynoic acid was the starter unit for polyketide extension.[16] Additionally, end-sequencing of these two fosmids revealed that the 3′-end of fosmid pHctI encoded a heterocyclization domain and the 5′-end of fosmid pHctII contained a putative halogenase related to BarB1 from the barbamide gene cluster.[18] Because, this genetic architecture was consistent with our predictions for hectochlorin biosynthesis, fosmids pHctI and pHctII were selected for further sequence analysis.

DNA Sequencing and Analysis of the Hectochlorin Biosynthetic Gene Cluster. The DNA sequences of the two overlapping fosmids, pHctI and pHctII, were determined by shotgun cloning and assembled into a 42 kb contiguous region that contained the 38 kb *hct* gene cluster (Figure 2A). Flanking the *hct* cluster at the 5′-end are three small ORFs (I, II, and III). ORF I encodes a 238 amino acid reverse transcriptase that is 58% identical to a similar protein from *Trichodesmium erythraeum*, ORF II encodes a putative endonuclease with 50% identity to a similar protein from *Anabaena variabilis*, and ORF III is similar to a hypothetical protein from *Nostoc punctiforme* PCC73102. At the 3′-end are two small ORFs (IV, V) that also encode hypothetical proteins from *T. erythraeum* and helped to confirm the outer boundaries of the hectochlorin biosynthetic gene cluster. The gene cluster exhibits a relatively low overall G+C content (~42%) that is comparable to the G+C content encountered in cyanobacterial genomes (http://www.kazusa.or.jp/cyano) and other secondary metabolite gene clusters from *L. majuscula*.[13,16,18,19]

The *hct* gene cluster consists of eight ORFs with the first ORF *hctA* encoding a protein that is similar to acyl-ACP synthetases (Table 1). Sequence analysis revealed a 59% identity to JamA in

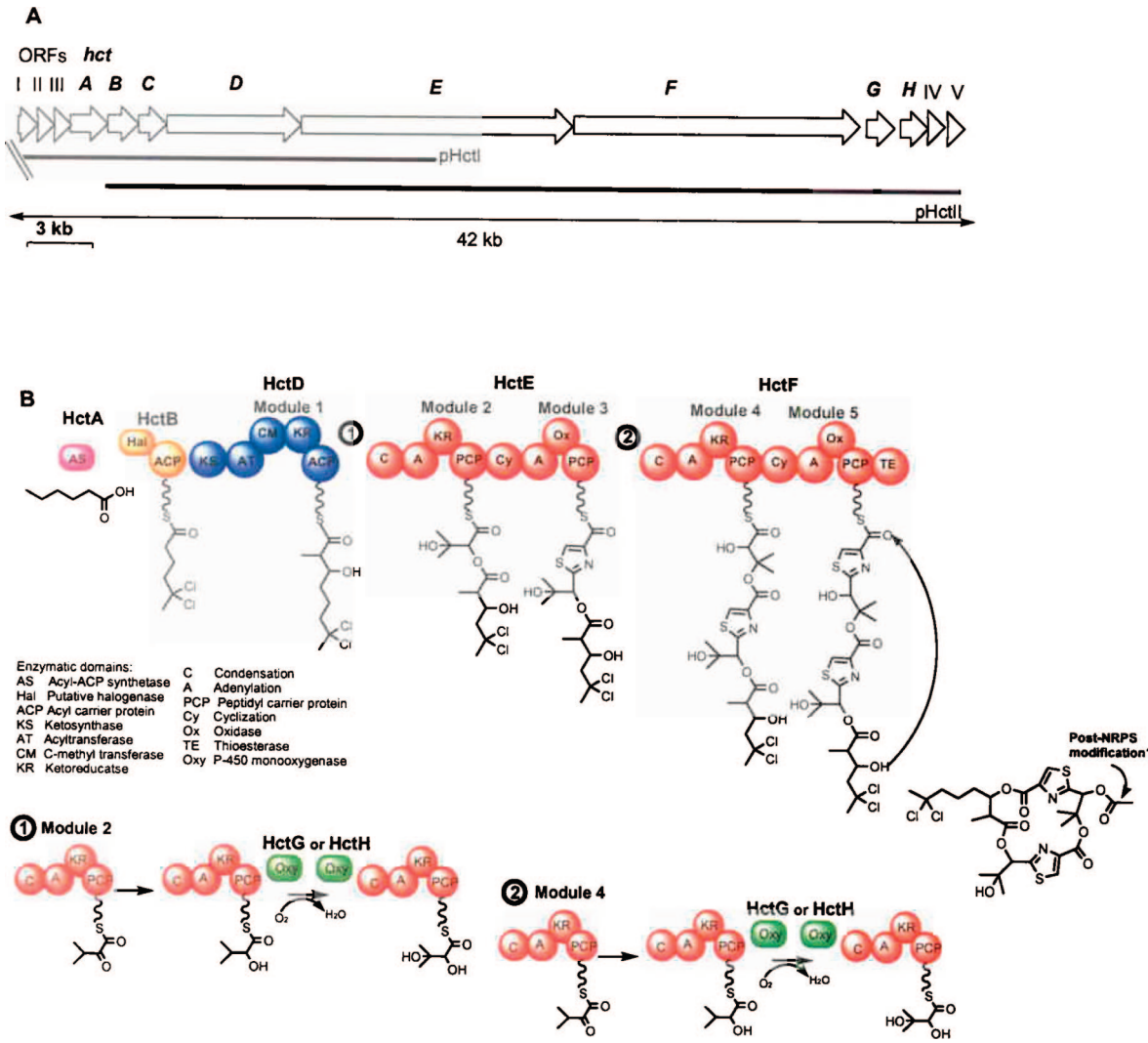

Figure 2. Biosynthetic gene cluster and proposed biosynthesis of hectochlorin (**1**). (A) Genetic map of the *hct* gene cluster spanning ~38 kb. Inserts from two overlapping fosmids, pHctI and pHctII, were sequenced to map the cluster. (B) Proposed biosynthesis of hectochlorin. Insets 1 and 2: In modules 2 and 4, the adenylation domains activate 2-oxo-isovaleric acid, which is reduced *in situ* to 2-hydroxyisovaleric acid by the embedded KR domain and further oxidized by HctG or HctH to form 2,3-dihydroxyisovaleric acid.

Table 1. Deduced Functions of the Proteins in the *hct* Biosynthetic Gene Cluster

protein	size (amino acids)	proposed function	sequence similarity (protein, organism)	identity/similarity	ref
ORF I	~238	reverse transcriptase	*Trichodesmium erythaeum*	58%, 73%	
ORF II	119	HNH endonuclease	*Anabaena variabilis ATCC29413*	50%, 65%	
ORF III	51	hypothetical protein	*Nostoc punctiforme ATCC73102*	60%, 73%	
HctA	606	acyl-ACP synthetase	JamA *L. majuscula* JHB	60%, 78%	18
Hct B	484	putative halogenase, acyl carrier protein	BarB1 *L. majuscula* 19L	29%, 47%	20
			JamC *L. majuscula* JHB	35%, 54%	18
HctC	408	transposase	TnpB *Helicobacter pylori*	33%, 55%	26
HctD	1924	polyketide synthase	*L. majuscula* JHB	56%, 72%	18
HctE	3356	nonribosomal peptide synthetase	BarE/BarG	46%, 63%	20
			L.majuscula 19L	57%, 71%	
HctF	3945	nonribosomal peptide synthetase	BarE/BarG	45%, 62%	20
			L.majuscula 19L	60%, 72%	
HctG	447	cytochrome P450, monooxygenase	TaH (P450 hydroxylase) *Myxococcus xanthus*	36%, 54%	33
			(P450 monooxygenase) *Polyangium cellulosum*	35%, 56%	spirangiene biosynthesis, Accession no. AJ505006
HctH	444	cytochrome P450, monooxygenase	(P450 monooxygenase) *Polyangium cellulosum*	36%, 57%	spirangiene biosynthesis, Accession no. AJ505006
			TaH (P450 hydroxylase *Myxococcus xanthus*	34%, 53%	33
ORF IV	90	hypothetical protein	*Trichodesmium erythraeum*	70%, 83%	
ORF V	222	Hypothetical protein	*Trichodesmium erythraeum*	58%, 75%	

the jamaicamide biosynthetic gene cluster.[16] Next, the product of *hctB* has two distinct domains. The N-terminal domain shows closest similarity to barbamide BarB1 (29% identity), coronamic acid CmaB (33% identity), and syringomycin SyrB2 (33% identity).[18,20–22] It also shows weak homology to Fe^{2+}/2-oxoglutarate-dependent hydroxylases such as phytanoyl Co-A hydroxylases (PAHX) and possesses all of the conserved amino acids necessary for cofactor binding (Figure 3).[23] The C-terminal domain is 35% identical to JamC, an acyl carrier protein in the jamaicamide gene cluster.[16]

The *hctC* gene is similar to insertion sequence (IS) elements from other bacteria and encodes for a putative transposase. BLAST analysis of the transposase showed that it was most similar (~31% identity) to TnpB encoded by an IS*605* type element from *Helicobacter pylori*[24] and transposases encoded by other IS elements such as IS*1341* from the thermophile PS3,[25] IS*891* from the cyanobacterium *Anabaena* M-131,[26] and IS*1136* from the erythromycin gene cluster in *Saccharopolyspora erythraea*.[27]

About 50 bp downstream from *hctC* is located *hctD*, which encodes a type I PKS module with KS-AT-CM-KR-ACP domain organization. The *hctD* gene shows highest similarity to PKS genes from the jamaicamide A and curacin A biosynthetic pathways, and all five domains show typical conserved core motifs when aligned with PKS genes from these two gene clusters (Figure 4).[13,16]

The stop site of *hctD* overlaps with the start site of the next ORF, *hctE*, which encodes for a bimodular NRPS. Initial analysis of the first module revealed the presence of a KR domain embedded between the A and PCP domains, which is rare in NRPS modules. The A domain is most similar to the A domain from BarE of the barbamide biosynthetic pathway (46% identity) and also shows similarity to A domains from the microcystin McyG and nodularin NdaC (41% identity).[18,28,29] A notable feature of the residues that contribute to substrate specificity in the A domain is substitution of the conserved aspartate residue (Asp235) involved in interactions with the amino group of the cognate amino acid by valine (Val235) at this position. BLAST analysis of the KR domain revealed it had weak similarity to KR domains from PKS (~23% identity), but higher similarity to a KR from the cereulide biosynthetic gene cluster (ces) (31% identity).[30] The second module in HctE contained all of the required domains for the adenylation and heterocyclization of cysteine and is most similar to BarG from the barbamide biosynthetic gene cluster (57% identity) and CurF from the curacin A biosynthetic pathway (58% identity, Figure 5)[13,18] This module also contains an FMN-dependent oxidase that is embedded between the A9 and A10 conserved motifs of the adenylation domain, which is likely involved in the formation of the thiazole ring.[31]

```
        121                                 150
BarB1   ...SKMNYDR HLDIQALSDI ISHPKIAHRV
 PAHX   ...MITKVQD FQEDKELFRY CTLPEILKYV
BarB2   ...NKMNYDR HLDIKALNDI ISHPKVVDRV
SyrB2   GT.NIANYDR HLDDDFLASH ICRPEICDRV
 CmaB   ....ITNYDR HLDVDLLTSH VFRKEIVHRV
 HctB   VLNLNTSRDL HLFHQPIAEM FKNNKIVRVL
        171                                 200
BarB1   GDEG.TDWHQ AESFVEFEGK SK...LVPTE
 PAHX   DSGKKTSRHP LHQDLHYFPF RP.....S..
BarB2   GDEG.TDWHQ AESFVEFEGE SK...LVPTE
SyrB2   GDEG.TDWHQ ADTFANASGK PQ...IIWPE
 CmaB   GNEG.TDWHQ ADTFAHAPAS RN...WCAR.
 HctB   GEGE.IKWHQ VYDSYDPSAY DPQKPALLYP
        301                                 330
BarB1   FTSRCMHGSE PNTSGSS.IR YGWSTRFVPT
 PAHX   FHPLLIHGS. .GQNKTQGFR KAISCHFASA
BarB2   FTSRCMHGSE PNTSESS.IR YGWSTRFVST
SyrB2   FWSTLMHASY PHSGESQEMR MGFASRYVPS
 CmaB   FGSMLMPSSL PNSTTDK... ....TAWAMS
 HctB   FTERVMHSSP PNKSKQQ.RR LGINGRYVPP
```

Figure 3. Alignment of barbamide BarB1 and B2, syringomycin SyrB2, coronamic acid CmaB, and PAHX with HctB. All enzymes share putative conserved residues (highlighted) involved in binding iron(II) and 2-oxoglutarate cofactors.

Next in the cluster is *hctF*, a bimodular NRPS whose domain architecture is identical to that of HctE. The predicted HctE and HctF proteins share a 76% overall identity. The *hctE* and *hctF* genes also share overlapping start and stop sites, suggesting that *hctDEF* are transcribed as an operon. A thioesterase domain is located at the C-terminal end of HctF and is proposed to catalyze hydrolysis from the biosynthetic assembly line and cyclization.

The next ORF, *hctG*, begins about 600 bp downstream from the end of *hctF* and encodes for a putative P450 monooxygenase. The *hctG* gene is followed by another ORF *hctH*, whose product also shows similarity to putative P450 monooxygenases. While HctG and HctH are most similar to each other (47% identity), they also show similarity to a P450 monooxygenase involved in spirangiene biosynthesis (accession number: AJ505006) (HctG 35% identity, HctH 36% identity) and TaH, a P450 hydroxylase from the TA biosynthetic gene cluster (HctG 36% identity, HctH 34% identity).[32]

Cloning, Expression, and Functional Analysis of Adenylation Domains from the Hectochlorin Biosynthetic Gene Cluster. In order to obtain biochemical confirmation of the predicted adenylation domain specificities, both A domains from HctE were overexpressed as 6× His-fusion constructs, purified, and analyzed using the ATP-PPi exchange assay. The adenylation domains were PCR amplified and cloned into the pET20b(+) expression vector. When *E. coli* BL21 (DE3) cells harboring these pET20b(+) constructs were grown at temperatures above 20 °C, only insoluble protein was obtained. However, soluble protein could be generated when cells were grown at 18 °C for 12–18 h with IPTG induction (Figure 6A). Western blot analysis using α-His antibodies was performed to confirm heterologous protein expression (data not shown).

The relative activities of the various amino acid substrates in the ATP-PPi exchange assay are displayed in Figures 6B and C. The first adenylation domain in HctE ($HctE_{IVA}$) was found to activate D-2-hydroxyisovaleric acid (D-2-HIV) to the greatest extent; however, it also activated 2-oxoisovaleric acid (O-IVA), and 2-oxo-3-hydroxyisovaleric acid to an appreciable level. $HctE_{IVA}$ was also shown to activate L-2-hydroxyisovaleric acid (L-2-HIV), both stereoisomers of DHIV, and leucine (Leu) to a smaller extent, while several other standard amino acid substrates were activated to a much less degree. These relative activities indicate that $HctE_{IVA}$ possesses fairly relaxed substrate specificity (Figure 6B). These results are consistent with the arrangement of amino acid residues in the binding pocket of the A domain in which the conserved Asp235 involved in the interaction with the amino group of the amino acid substrate is replaced by Val235, a substitution typical for a keto-acid or hydroxy-acid substrate.[33] In the case of the second adenylation domain ($HctE_{Cys}$), L-cysteine was activated to the highest level (=100%) and L-serine was activated up to 40%, while activation of L-glycine, L-tryptophan, and L-histidine were equivalent to background (Figure 6C).

Biosynthesis of Hectochlorin. This report describes the cloning and preliminary characterization of 38 kb of DNA from the marine cyanobacterium *Lyngbya majuscula* that is putatively involved in the biosynthesis of hectochlorin (**1**). The overall genetic architecture and domain organization encountered in this cluster is consistent with the arrangement expected for the biosynthesis of hectochlorin, and its identity is further strengthened by biochemical characterization of the adenylation domains.

HctA, an acyl-ACP synthetase homologue, is proposed to activate free hexanoic acid in an ATP-dependent manner and thus initiate hectochlorin biosynthesis. HctA shares a high similarity to jamaicamide JamA (obtained from the same *L. majuscula* strain), which has been shown to biochemically activate 5-hexynoic acid through

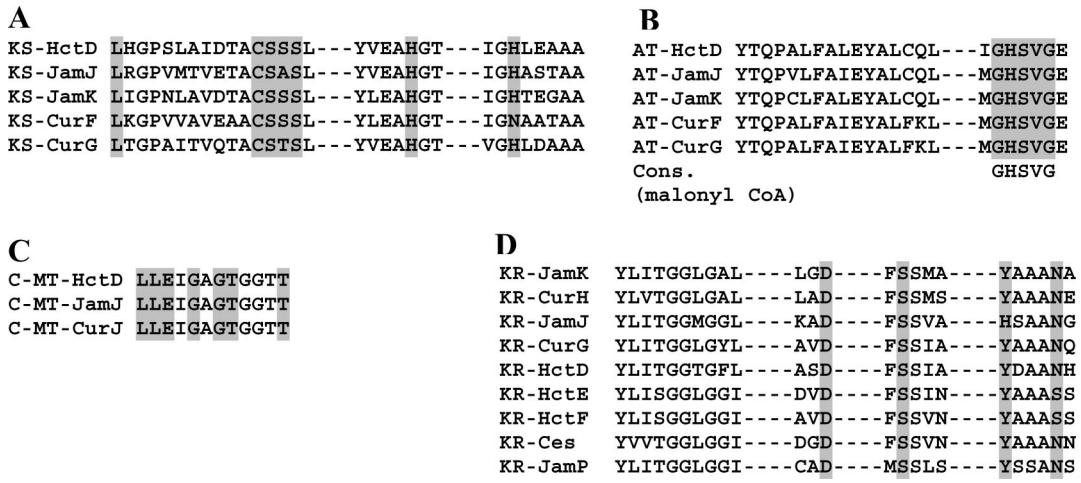

Figure 4. Sequence alignments of key motifs deduced in the PKS domains of HctD with motifs from other cyanobacterial PKSs (jamaicamide A (**5**) and curacin A PKS genes). KS domain alignment (A), AT domain specific for malonyl-CoA (B), alignment of C-methyl transferase (C), and alignment of KR domains from jamaicamide A, curacin A, cereulide, and hectochlorin biosynthetic gene clusters (D).

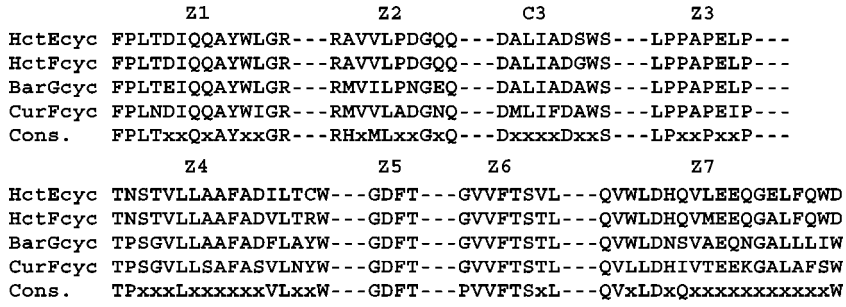

Figure 5. Alignment of heterocyclization domains from barbamide BarG and curacin A CurF with heterocyclization domains from the *hct* gene cluster showing the presence of consensus core motifs.

an ATP-PPi exchange assay.[16] The acyl group is then likely transferred from the acyl-adenylate to the ACP domain in the C-terminal region of HctB. The N-terminal region of HctB shows similarity to barbamide BarB1 and B2,[18] coronamic acid CmaB,[20] and syringomycin SyrB2 and weak similarity to phytanoyl-CoA hydroxylases.[22,23] Sequence alignments of these proteins with HctB show that they share conserved residues involved in cofactor binding, suggesting that the enzyme mechanism requires iron(II) and 2-oxoglutarate as cofactors (Figure 3).[23,34] Using isotope-labeled precursor feeding experiments and *in vitro* protein biochemical function assays, BarB1 and B2 were shown to mediate the transformation of leucine to trichloroleucine through tandem reactions of these two non-heme Fe^{II} halogenases during the biosynthesis of barbamide.[35,36] Using similar protein biochemical assays, SyrB2 was shown to be involved in the chlorination of threonine during syringomycin biosynthesis and CmaB converts isoleucine to coronamic acid through a cryptic chlorination reaction in coronatine biosynthesis.[37,38] Hence, we speculate that HctB is involved in the formation of the gem-dichloro group in hectochlorin on the basis of its location in the cluster and proposed biochemical properties.

Located between *hctB* and *hctD* is an IS element that encodes for a putative transposase. IS elements are fairly small segments (1.5–2.0 kb) with a simple genetic organization that encode functions involved in their mobility.[39] IS elements are believed to play a major role in the plasticity of bacterial genomes and are often involved in the assembly of gene clusters with specialized functions such as pathogenicity, antibacterial activity, or catabolic pathways.[40] The occurrence of transposases at the boundaries of several cyanobacterial gene clusters has been previously reported;[16,28,29,41,42] however, it is interesting to note that the HctC transposase is present within the putative *hct* gene cluster. It is likely that these transposases may play an important role in the acquisition and distribution of secondary metabolite gene clusters between different strains and genera.

The chlorinated hexanoate initiator of hectochlorin biosynthesis is next proposed to undergo one round of ketide extension by the PKS module (KS-AT-CM-KR-ACP) found in HctD. The AT domain in this module appears to be specific for malonyl-CoA (Figure 4B).[43] The C-MT transferase possesses a signature motif (LExGxGxG) that is also present in the C-methyl transferases from the jamaicamide A and curacin A biosynthetic gene clusters (Figure 4C). This is consistent with precursor feeding experiments to several cyanobacterial polyketide pathways in which methyl branching at C-2 positions all derive from SAM.[3,13,16]

Next, HctE is composed of two NRPS modules that are proposed to incorporate DHIV and cysteine. Module 2, which presumably catalyzes the incorporation of DHIV, is unusual in that it possesses a KR domain embedded between the A and PCP domains. Analysis of the A-domain binding pocket residues that confer substrate specificity using the method of Challis et al. revealed that the conserved Asp235 involved in ionic interaction with the amino group of the substrate amino acid was replaced by Val235 (Table 2).[44] This type of codon specificity suggests a preference for a non α-amino acid substrate such as a hydroxy- or keto-acid. Furthermore, this A domain is most closely related to the barbamide BarE A domain, which also lacks the conserved Asp235 and shows high specificity for both the nonchlorinated and trichlorinated 2-oxo derivatives of L-leucine and only modest activation of corresponding α-amino acids.[18,35] Other previously characterized cyanobacterial A domains with the missing Asp235 residue include the McyG and NdaC A domains, both of which are involved in activating a phenylacetate starter unit in the biosynthesis of microcystin and nodularin, respectively (Table 2).[28,29]

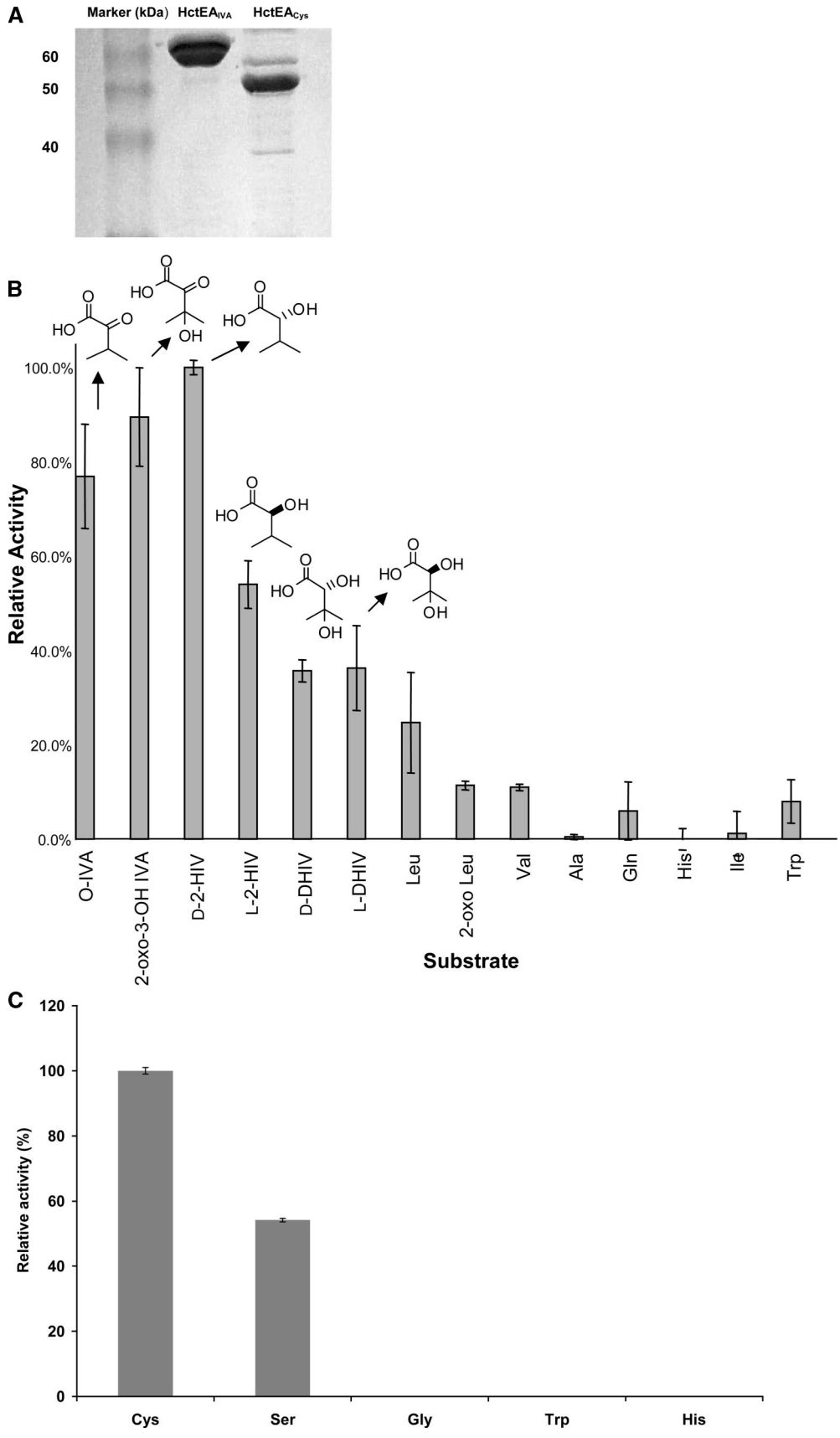

Figure 6. Expression of adenylation domains and substrate-dependent ATP-PPi exchange activity. (A) HctE$_{IVA}$ and HctE$_{Cys}$ adenylation domains were overexpressed as 6× His-fusion proteins and purified. Results are shown for the ATP-PPi exhange assay for HctE$_{IVA}$ (B) and HctE$_{Cys}$ (C) and represent average of duplicate experiments. Abbreviations: O-IVA, 2-oxoisovaleric acid; 2-HIV, 2-hydroxyisovaleric acid; 2-oxo-3-OH IVA, 2-oxo-3-hydroxyisovaleric acid; DHIV, 2,3-dihydroxyisovaleric acid; Val, L-valine; 2-oxo-Leu, 2-oxo L-leucine; Leu, L-leucine; Ala, L-alanine; His, L-histidine; Gln, L-glutamine; Trp, L-tryptophan; Cys, L-cysteine; Ser, L-serine; Gly, L-glycine.

Table 2. Binding Pocket Amino Acids Involved in Substrate Recognition Determined Using the Method of A Domains from Module 2 ($HctE_{IVA}$) and Module 4 ($HctF_{IVA}$) (this study) Were Compared to Residues from Previously Characterized A Domains BarE, McyG, and NdaC Involved in Activating Non Amino Acid Substrates[a]

	binding pocket amino acids								
A domain	235	236	239	278	299	301	322	330	substrate
$HctE_{O\text{-}IVA}$	V	G	V	W	L	A	L	F	2-oxoisovaleric acid
$HctF_{O\text{-}IVA}$	V	G	V	W	L	A	L	F	2-oxoisovaleric acid
BarE	V	G	I	I	V	G	G	T	2-oxoleucic acid
NdaC	V	G	I	W	V	A	A	S	phenylacetate
McyG	V	G	I	W	V	A	A	S	phenylacetate
GrsA-Phe	D	A	W	T	I	A	A	I	phenylalanine

[a] The binding pocket amino acid residues correspond to GrsA-Phe A domain numbering.[41]

Biochemical support for $HctE_{IVA}$ function was obtained through the ATP-PPi exchange assay in which the A domain was found to activate the 2-oxo, 2-oxo-3-hydroxy, and D-2-hydroxyisovaleric derivatives preferentially over other substrates (Figure 6B). However, the exchange assay also shows that this A domain activates L-2-HIV, DHIV, and Leu, to a lesser degree. The presence of a KR domain in an NRPS module has been reported in the cereulide,[30] valinomycin,[45] and cryptophycin biosynthetic gene clusters.[46] In cereulide biosynthesis, the function of the KR was validated by labeled precursor feeding studies, which concluded that L-valine is transaminated to 2-oxoisovaleric acid (O-IVA) and then reduced to 2-OH IVA during the biosynthetic process.[47] More recently, this was biochemically demonstrated by showing the release of [^{14}C]-labeled α-hydroxy products from incubation of [^{14}C]-α-keto acyl-*S*-PCPs in the presence of NADPH.[48] Indeed, while the HctE and Ces KR domains exhibit only weak similarity to KR domains from PKS modules, they possess all of the key conserved amino acid residues required for enzymatic catalysis (Figure 4D).[49]

The biochemical results for $HctE_{IVA}$ contrast with the work of Magarvey et al., who investigated the embedded KR in the Ces pathway; in this latter pathway the two KRs showed a strong substrate preference for α-keto acids over α-hydroxy acids. The only hydroxy acid activated to an appreciable degree in the cereulide pathway was D-HIV by CesB;[48] this is interesting because the final product contains only L-HIV and indicates that the adenylation domains for hydroxy acids must work in tandem with other components of the biosynthetic pathway (e.g., substrate pools, downstream processing) in order to incorporate the appropriate substrate with the correct configuration in the final product. Nevertheless, by similar experimental methods we have shown that $HctE_{IVA}$ has a more relaxed substrate specificity than CesA or CesB and may be capable of activating multiple substrates. The adenylation domain responsible for the incorporation of the hydroxy acid into cryptophycin also appears to accept multiple substrates due to the identification of cryptophycin analogues at this position.[46] However, activation of alternative substrates is likely not translated into alternative final products in hectochlorin; indeed, epimeric hectochlorin analogues were not isolated from the extract of the cultured organism. Ultimately, incubation studies using isotope-labeled precursors will be necessary to unequivocally determine the potential of this pathway to incorporate alternate configurations at these two DHIV sites.

Next, O-IVA undergoes β-hydroxylation, a reaction presumably catalyzed by either HctG or HctH. HctG and HctH are heme-dependent cytochrome P450 monooxygenases with conserved C-terminal motifs, GxHxCxGxxLxR, which are involved in heme-binding and show greatest similarity to TaH from the TA antibiotic biosynthetic gene cluster and a P450 monooxygenase from spirangiene biosynthesis (accession number: AJ505006).[50] TaH is responsible for a specific non-PKS hydroxylation at C-20 of antibiotic TA that occurs post-assembly.[32] P450 monooxygenases involved in post-assembly modification are also seen in other polyketide gene clusters such as those involved in the biosynthesis of erythromycin and epothilone.[51,52] However, P450 monooxygenases that function during assembly to generate β-hydroxy amino acid substrates are encountered in NRPS gene clusters such as those involved in the biosynthesis of certain peptide antibiotics. A common theme in these latter gene clusters is the occurrence of a heme oxygenase in conjunction with an A-PCP didomain. The substrate amino acid is activated by the A domain and then transferred to the holo-PCP domain to yield an aminoacyl-*S*-enzyme intermediate that is β-hydroxylated by the heme monooxygenase. In novobiocin biosynthesis, NovH (A-PCP) activates the amino acid tyrosine to L-tyrosyl-AMP, which accumulates on its PCP domain. Its partner enzyme, NovI, is a heme-dependent cytochrome P450 monooxygenase that specificially hydroxylates this intermediate to form β-hydroxytyrosine.[53] Similarly, during biosynthesis of the nucleoside antibiotic nikkomycin, NikP1 activates histidine and the L-histidinyl-*S*-NikP1 serves as a substrate for the heme monooxygenase NikQ, which introduces the β-hydroxy group.[54] Such heme monooxygenases are thought to have evolved in conjunction with corresponding NRPSs to act on substrates that are presented as L-aminoacyl-*S*-enzyme complexes, thus ensuring that only a select pool of the substrate is modified.[55] Although the HctG and HctH are more similar to monooxygenases that catalyze post-assembly hydroxylations, it is plausible that they mechanistically resemble NovI and NikQ by catalyzing the β-hydroxylation of O-IVA in *trans* position during the assembly process to generate 2,3-dihydroxyisovaleric acid units (Figure 2B). It is interesting to note that there are two heme-dependent oxygenases in the *hct* gene cluster; however, as discussed above, the extent of their involvement, timing of β-hydroxylation, and biochemical mechanism await further characterization.

Table 3. Amino Acid Residues Involved in Substrate Recognition from Adenylation Domains Module 3 ($HctE_{Cys}$) and Module 5 ($HctF_{Cys}$) Were Identified Using the Method of A Domains and Compared to Other Cyanobacterial A Domains Involved in Activating Cysteine $BarGA_2$ and CurF[a]

	binding pocket amino acids								
A domain	235	236	239	278	299	301	322	330	predicted substrate
$HctE_{Cys}$	D	L	Y	N	L	S	L	I	cysteine
$HctF_{Cys}$	D	L	Y	N	A	L	L	I	cysteine
$BarG_{Cys}$	D	L	Y	N	L	S	L	I	cysteine
$CurF_{Cys}$	D	L	Y	N	V	D	L	I	cysteine

[a] The binding pocket amino acid residues correspond to GrsA-Phe A-domain numbering.[41]

Module 3 shows high similarity to NRPSs specific for the incorporation and heterocyclization of cysteine. Analyses of the amino acids in the binding pocket of the adenylation domain are consistent with conserved amino acid residues for cysteine recognition (Table 3). Notably, the condensation domain in this module shows the presence of conserved core motifs required for heterocyclization (Figure 5).[56] *In vitro* biochemical characterization of the $HctE_{Cys}$ domain demonstrated that it activates L-cysteine preferentially (Figure 6C). An FMN-dependent oxidase domain is embedded between the A9−A10 core motifs of the adenylation domain and presumably catalyzes the oxidation of the thiazoline to a thiazole. Similar oxidase domains involved in thiazole ring formation from the epothilone and bleomycin gene clusters have been biochemically characterized and show high similarity to this domain from the *hct* gene cluster (48% identity).[57]

Next, modules 4 and 5 in HctF are believed to catalyze the incorporation of the second DHIV and cysteine units, respectively. The modular organization of HctF is identical to that present in HctE. The A domain in module 4 exhibits a remarkably high

similarity to the A domain in module 2 (~95% identical) with both adenylation domains sharing identical binding pocket amino acid residues (Table 2). Thus, the A domain in module 4 is expected to catalyze the activation and incorporation of the second O-IVA that is reduced *in situ* to 2-HIV by the embedded KR domain and then further oxidized by HctG or HctH to form DHIV. Although the incorporation of the two DHIV units is catalyzed by modules that have the same domain organizations, it is intriguing that their regiochemical orientation is different, with the secondary alcohol of DHIV-1 forming the ester with the preceding polyketide unit and the tertiary alcohol of DHIV-2 forming the ester with the preceding cysteine-derived residue. It is unknown how the alternate orientations of DHIV are achieved by these very similar modules.

Module 5 in HctF appears to catalyze the incorporation of a second L-cysteine unit into hectochlorin (**1**). The A-domain binding pocket amino acids are in consensus for the recognition of cysteine (Table 3).[44] Conserved regions were again identified for the heterocyclization domain that catalyzes the cyclization of cysteine (Figure 5). An oxidase domain is present between the A9 and A10 conserved motifs and is again likely involved in the oxidation of thiazoline to thiazole. Finally, an embedded type I thioesterase that possesses the conserved GxSxG motif is present at the C-terminal end of HctF and may be involved in catalyzing product release and cyclization. A candidate gene encoding a protein that introduces the acetate group at C-14 in hectochlorin (**1**) is lacking and leads us to speculate that this may be catalyzed by an acetyl transferase not present in the *hct* gene cluster and most likely occurs as a post-NRPS modification.

Conclusions

The natural products chemistry of marine cyanobacteria is amazingly diverse and is a direct reflection of the unique biosynthetic capabilities of these organisms. *L. majuscula* is clearly the most prominent producer of novel secondary metabolites among these organisms. The *hct* gene cluster adds to the growing repertoire of cyanobacterial hybrid PKS-NRPS pathways and contains several rare elements. An interesting feature of this cluster is the juxtaposition of two multimodule ORFs with identical organization, suggesting that it may have arisen through a gene duplication event. Sequence alignment of HctE and HctF indicates >90% identity in certain regions.

Analysis of the putative *hct* gene cluster has revealed rare motifs such as the presence of an embedded KR domain in an NRPS module, which is likely involved in reducing an O-IVA precursor. This is only the second report of the presence of a KR domain in an NRPS module from a cyanobacterial gene cluster, to the best of our knowledge. Other notable features include an acyl-ACP synthetase, a putative halogenase, two P450 monooxygenases, and the presence of an IS element within the cluster. Another striking feature of this gene cluster is the presence of homologues of recognizable genetic elements such as JamA from the jamaicamide A gene cluster and BarB1 from the barbamide gene cluster that are found in new combinations (HctA and HctB).

The isolation of each new cyanobacterial natural product biosynthetic gene cluster has illustrated the remarkable versatility of cyanobacterial biosynthetic machinery and provides the foundation for harnessing the activity of unique tailoring enzymes. The continued discovery of new molecules from these metabolically talented organisms will provide a bountiful source of novel bioactive natural products that may be developed as potential therapeutics. Additionally, efforts focused at dissecting the biosynthesis of these metabolites at the molecular genetic level along with whole-genome sequencing of several cyanobacteria will enable full exploitation of the biochemical diversity of these organisms.

Experimental Section

Bacterial Strains and Culture Conditions. *Escherichia coli* strain DH10B was routinely used in this study as a host for DNA cloning and sequencing, while *E. coli* strain EPI 300 was used as a host for construction of the fosmid library and amplification of fosmid clones. *E. coli* DH10B containing pGEM-T Easy, pBluescript, or pGEM3ZF(+) was grown at 37 °C overnight in Luria–Bertani medium (LB) with a final concentration of ampicillin at 100 μg/mL. *E. coli* EPI 300 harboring fosmid clones were grown at 37 °C in LB medium with chloramphenicol at a final concentration of 12 μg/mL. *Lyngbya majuscula* JHB was collected from Hector Bay in 1996 and is maintained as a unialgal culture in our laboratory.[7]

DNA Isolation and Construction of Fosmid Library. DNA from *L. majuscula* JHB for PCR amplification was obtained using DNeasy Plant Mini Kit (Qiagen, Valencia, CA). High molecular weight DNA for construction of the library from *L. majuscula* JHB was extracted using phenol followed by purification with the Genomic Tip Kit (Qiagen, Valencia, CA). The high molecular weight DNA was size selected (~40 kb), end repaired, and ligated into the copy control fosmid vector pCC1FOS according to instructions provided with the Copy-Control Fosmid Library Production Kit (Epicenter, Madison, WI). Plasmid DNA isolation was carried out using commercially available kits (Qiagen, Valencia, CA). Other standard DNA manipulations such as restriction digests, ligations, and transformations were carried out using standard methods.[58] Library screening, labeling of probes, detection, and Southern analysis were performed according to the protocols provided with the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche Molecular Biochemicals, Manheim, Germany).

PCR Cloning of Heterocyclization−Adenylation Domain Probe. PCR amplification of probe fragments used to screen the genomic library in this study was carried out using *Taq* DNA-polymerase (Invitrogen, Carlsbad, CA) and the manufacturer's recommendation regarding template and primer concentrations in an Eppendorf Mastercycler. The following thermocycler conditions were used for amplification: denaturation at 94 °C for 30 s, annealing at 48 °C for 30 s, extension at 72 °C for 1 min for 30 cycles. Degenerate primers were designed on the basis of conserved domains found in the heterocyclization−adenylation domain regions (HetForward-5′-CCG GAT AAG CGA CAT GAA CCN TTY CCN YTN A-3′ and AdenReverse-5′-CAT CAG GGC GGG CAC NSW RTT CCA-3′) (Supporting Information). These primers were used to amplify an approximately 1.7 kb fragment. The PCR product was purified using a commercially available kit (Qiagen, Valencia, CA) and cloned into pGEM-T Easy (Promega, Madison, WI), and unique clones were sequenced. Amplification of an approximately 700 bp region of the ketosynthase domain was carried out using previously designed primers KS1Up 5′-MGI GAR GCI HWI SMI ATG GAY CCI CAR CAI MG-3′ and KSD1 5′-GGR TCI CCI ARI SWI GTI CCI GTI CCR TG-3′.[17] All primers were obtained from Invitrogen (Carlsbad, CA).

Sequence Analysis of Fosmids. Sequencing of fosmids pHctI and pHctII was carried out by shotgun cloning. Fosmid DNA was subjected to partial digest using *Alu*I or *Sau*3AI (0.10–0.25 units/μg of fosmid DNA for 1 h). The partially digested fosmid DNA was then separated on a 1% agarose gel and fragments of size 1–2.5 kb were selected. The *Alu*I fragments were ligated into pGEM3Zf-(+) treated with *Sma*I/CIAP (calf intestinal alkaline phosphatase), *Sau*3AI fragments were cloned into pBluescript that was digested with *Bam*HI and CIAP-treated, and random subclones were selected for sequencing reactions. Sequencing was carried out using Big Dye III terminator cycle sequencing (PE biosystems) at the Central Services Laboratory, Center for Gene Research and Biotechnology, Oregon State University. Resulting gaps in the sequences were filled using primer walking with sequence-specific primers. The sequences were edited and assembled using VectorNTI's ContigAssembly software program (InforMax Inc., Frederick, MD). Analyses of DNA and protein sequences were carried out using the NCBI (National Center for Biotechnology Information) BLAST server. All alignments were conducted using ClustalW.[59]

Cloning, Expression, and Purification of Adenylation Domains. The adenylation domains of HctE were amplified by PCR using Vent DNA polymerase (New England Biolabs) according to manufacturer's recommendations. The following primers (5′ tail in italics; restriction digest sites in bold): $HctE_{IVA}$-Forward3–5′*GGA ATT C***CA TAT G**GA AAC TGA CAA CTT GAT CGA C 3′ and $HctE_{IVA}$-Reverse3–5′ *CCG* **CTC GAG** GGA AGA CAG ATT AAA AAC AGG 3′ were used to amplify the adenylation domain from the first module in HctE (amino acids 787−1391) and the primers $HctE_{Cys}$-Forward3–5′*GGA ATT C***CA**

TAT GTT CCA GAG TTA CAC AGA TTA T 3′ and $HctE_{Cys}$-Reverse3–5′ *CCG***CTC GAG** ATC AAA TTC TGA CAG TGG TAA were used to amplify the adenylation domain $HctE_{Cys}$ (amino acids 2450–2956). The PCR products were cloned in-frame with a carboxy-terminal 6xHis fusion at the *Nde*I and *Xho*I sites of the pET20b(+) expression vector (Novagen). Both constructs were verified by sequencing. Overexpression of proteins was carried out in *E. coli* BL21 (DE3), and both proteins were purified in an identical manner. Overnight cultures expressing $HctE_{IVA}$ A domain or $HctE_{Cys}$ A domain were diluted 1:100 in 1 L of LB broth containing 100 μg/mL ampicillin and grown at 37 °C for 2 h prior to IPTG induction and then grown 12–18 h at 18 °C with shaking (150 rpm). Cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris-HCl [pH = 8], 300 mM NaCl, and 20 mM imidazole), and lysed by sonication for six 10 s bursts at 75 mAmps using a UPC 2000U sonicator (Ultrasonic Power Corporation, Freeport, IL). Recombinant proteins were purified using nickel chelate chromatography (Qiagen, Valencia, CA). Protein lysates were incubated with Ni-agarose beads for 2 h at 4 °C. The protein−Ni-agarose slurry was then washed three times with wash buffer (20 mM Tris-HCl [pH = 8], 300 mM NaCl, and 40 mM imidazole), and proteins were eluted with elution buffer (20 mM Tri-HCl [pH = 8], 300 mM NaCl, and 250 mM imidazole). Purified proteins were dialyzed overnight at 4 °C against storage buffer (50 mM Tris-HCl [pH = 8], 1 mM EDTA [pH = 8], 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, and 10% glycerol) using a Spectra/Por dialysis membrane with a molecular weight cutoff of 12 000–14 000 Da (Spectrum Laboratories Inc., Rancho Dominguez, CA). The dialyzed proteins were aliquoted, flash frozen in liquid nitrogen, and stored at −80 °C. Protein concentration was determined by Bradford protein assay, and protein purity was assessed by analysis of protein lysates on a 12% SDS-PAGE followed by staining with Coomassie blue staining. Identification of a His-tagged protein was confirmed through Western blot analysis using anti-His (C-term) antibody (Invitrogen).

Substrate-Dependent ATP-Pyrophosphate [^{32}PPi] Exchange Assay. All amino acid substrates used in this assay were obtained from Sigma (St. Louis, MO) or Lancaster (Windham, NH). 2-Oxo-3-hydroxyisovaleric acid and DHIV were synthesized using previously described methods.[9,60] The substrate-dependent ATP-PPi exchange assay was performed by incubating 2 mM of $HctE_{IVA}$ or $HctE_{Cys}$ in reaction buffer (50 mM Tris-HCl [pH = 8], 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT) containing 0.001 or 0.5 μCi tetrasodium [^{32}P] pyrophosphate, respectively (NEN-Perkin-Elmer, Boston, MA) and 2 mM of the respective amino acid or oxo-derivative for 1 h at 28 °C. The reaction was quenched by the addition of stop-mix (500 μL of 1.2% (w/v) activated charcoal, 0.1 M tetrasodium pyrophosphate, and 0.35 M $HClO_4$). The charcoal pellet was pelleted and washed three times with wash buffer (0.1 M tetrasodium pyrophosphate and 0.35 M $HClO_4$) and resuspended in 500 μL of H_2O. The bound radioactivity was determined using a Packard TriCarb 2900 liquid scintillation counter.

Accession Number. The nucleotide sequence of the hectochlorin biosynthetic gene described in this work has been submitted to GenBank under the accession number AY974560.

Acknowledgment. We thank D. J. Edwards for invaluable advice during sequence analysis of this cluster and critical reading of the manuscript; M. Musafija-Girt for assistance with laboratory culture of this strain; and T. Suyama and B. Han for providing 2-oxo-3-hydroxyisovaleric acid and 2,3-dihydroxyisovaleric acid, respectively. We also wish to acknowledge the Central Services Lab (Center for Gene Research and Biotechnology) at OSU for sequencing services. This research was supported by NIH grant CA83155.

Supporting Information Available: Design of heterocyclization−adenylation domain gene probes. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Newman, D. J.; Cragg, G. M.; Snader, K. M. *J. Nat. Prod.* **2003**, *66*, 1022–1037.
(2) Burja, A. M.; Banaigs, B.; Abou-Mansour, E.; Burgess, J. G.; Wright, P. C. *Tetrahedron* **2001**, *57*, 9347–9377.
(3) Gerwick, W. H.; Tan, L.; Sitachitta, N. *The Alkaloids*; Academic Press: San Diego, 2001; Vol. 57, pp 75–184.
(4) Van Wagoner, R. M.; Drummond, A. K.; Wright, J. L. C. *Adv. Appl. Microbiol.* **2007**, *61*, 89–213.
(5) Finking, R.; Marahiel, M. *Annu. Rev. Microbiol.* **2004**, *58*, 453–488.
(6) Weissman, K. J. *Philos. Transact. A Math. Phys. Eng. Sci.* **2004**, *362*, 2671–2690.
(7) Marquez, B. L.; Watts, K. S.; Yokochi, A.; Roberts, M. A.; Verdier-Pinard, P.; Jimenez, J. I.; Hamel, E.; Scheuer, P. J.; Gerwick, W. H. *J. Nat. Prod.* **2002**, *65*, 866–871.
(8) Sone, H.; Kondo, T.; Kiruyu, M.; Ishiwata, H.; Ojika, M.; Yamada, K. *J. Org. Chem.* **1995**, *60*, 4774–4781.
(9) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Mooberry, S. L. *J. Nat. Prod.* **2000**, *63*, 611–615.
(10) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2000**, *63*, 1437–1439.
(11) Milligan, K. E.; Marquez, B. L.; Williamson, R. T.; Gerwick, W. H. *J. Nat. Prod.* **2000**, *63*, 1440–1443.
(12) Cetusic, J. R. P.; Green, F. R.; Graupner, P. R.; Oliver, M. P. *Org. Lett.* **2002**, *4*, 1307–1310.
(13) Chang, Z.; SitachittaN.; Rossi, J. V.; Roberts, M. A.; FlattP. M.; Jia, J.; Sherman, D. H.; Gerwick, W. H. *J. Nat. Prod.* **2004**, *67*, 1356–1367.
(14) Williamson, R. T.; Sitachitta, N.; Gerwick, W. H. *Tetrahedron* **1999**, *40*, 5175–5178.
(15) Neilan, B. A.; Dittmann, E.; Rouhiainen, L.; Bass, R. A.; Schaub, V.; Sivonen, K.; Borner, T. *J. Bacteriol.* **1999**, *181*, 4089–4097.
(16) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817–833.
(17) Beyer, S.; Kunze, B.; Silakowski, B.; Muller, R. *Biochim. Biophys. Acta* **1999**, *1445*, 185–195.
(18) Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V. A.; Willis, C. L.; Sherman, D. H. *Gene* **2002**, *296*, 235–247.
(19) Edwards, D. J.; Gerwick, W. H. *J. Am. Chem. Soc.* **2004**, *126*, 11432–11433.
(20) Ullrich, M.; Bender, C. L. *J. Bacteriol.* **1994**, *176*, 7574–7586.
(21) Guenzi, E.; Galli, G.; Grgurina, I.; Gross, D. C.; Grandi, G. *J. Biol. Chem.* **1998**, *273*, 32857–32863.
(22) Zhang, J. H.; Quigley, N. B.; Gross, D. C. *J. Bacteriol.* **1995**, *177*, 4009–4020.
(23) Mukherji, M.; Chien, W.; Kershaw, N. J.; Clifton, I. J.; Schofield, C. J.; Wierzbicki, A. S.; Lloyd, M. D. *Hum. Mol. Genet.* **2001**, *10*, 1971–1982.
(24) Censini, S.; Lange, C.; Xiang, Z.; Crabtree, J. E.; Ghiara, P.; Borodovsky, M.; Rappuoli, R.; Covacci, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14648–14653.
(25) Murai, N.; Kamata, H.; Nagashima, Y.; Yagisawa, H.; Hirata, H. *Gene* **1995**, *163*, 103–107.
(26) Bancroft, I.; Wolk, C. P. *J. Bacteriol.* **1989**, *171*, 5949–5954.
(27) Donadio, S.; Staver, M. J. *Gene* **1993**, *126*, 147–151.
(28) Moffitt, M. C.; Neilan, B. A. *Appl. Environ. Microbiol.* **2004**, *70*, 6353–6362.
(29) Tillett, D.; Dittmann, E.; Erhard, M.; von Dohren, H.; Borner, T.; Neilan, B. A. *Chem. Biol.* **2000**, *7*, 753–764.
(30) Ehling-Schulz, M.; Vukov, N.; Schulz, A.; Shaheen, R.; Andersson, M.; Martlbauer, E.; Scherer, S. *Appl. Environ. Microbiol.* **2005**, *71*, 105–113.
(31) Du, L.; Chen, M.; Sanchez, C.; Shen, B *FEMS Microbiol. Lett.* **2000**, *189*, 171–175.
(32) Paitan, Y.; Orr, E.; Ron, E. Z.; Rosenberg, E. *Gene* **1999**, *228*, 147–153.
(33) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493–505.
(34) Kershaw, N. J.; Mukherji, M.; MacKinnon, C. H.; Claridge, T. D.; Odell, B.; Wierzbicki, A. S.; Lloyd, M. D.; Schofield, C. J. *Bioorg. Med. Chem.* **2001**, *11*, 2545–2548.
(35) Flatt, P. M.; O'Connell, S. J.; McPhail, K. L.; Zeller, G.; Willis, C. L.; Sherman, D. H.; Gerwick, W. H. *J. Nat. Prod.* **2006**, *69*, 938–944.
(36) Galonić, D. P.; Vaillancourt, F. H.; Walsh, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 3900–3901.
(37) Vaillancourt, F. H.; Yin, J.; Walsh, C. T. *Proc. Natl. Acad. Sci.* **2005**, *102*, 10111–10116.
(38) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'Connor, S. E.; Walsh, C. T. *Nature* **2005**, *436*, 1191–1194.
(39) Mahillon, J.; Chandler, M. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 725–774.
(40) Mahillon, J.; Leonard, C.; Chandler, M. *Res. Microbiol.* **1999**, *150*, 675–687.
(41) Becker, J. E.; Moore, R. E.; Moore, B. S. *Gene* **2004**, *325*, 35–42.
(42) Christiansen, G.; Fastner, J.; Erhard, M.; Borner, T.; Dittmann, E. *J. Bacteriol.* **2003**, *185*, 564–72.
(43) Haydock, S. F.; Aparicio, J. F.; Molnar, I.; Schwecke, T.; Khaw, L. E.; Konig, A.; Marsden, A. F.; Galloway, I. S.; Staunton, J.; Leadlay, P. F. *FEBS Lett.* **1995**, *374*, 246–248.

(44) Challis, G. L.; Ravel, J.; Townsend, C. A. *Chem. Biol.* **2000**, *7*, 211–224.
(45) Cheng, Y. *ChemBioChem* **2006**, *7*, 471–477.
(46) Magarvey, N. A.; Beck, Z. Q.; Golakoti, T.; Ding, Y.; Huber, U.; Hemschidt, T. K.; Abelson, D.; Moore, R. E.; Sherman, D. H. *ACS Chem. Biol.* **2006**, *1*, 766–779.
(47) Kuse, M.; Franz, T.; Koga, K.; Suwan, S.; Isobe, M.; Agata, N.; Ohta, M. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 735–739.
(48) Magarvey, N. A.; Ehling-Schulz, M.; Walsh, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 10698–10699.
(49) Reid, R.; Piagentini, M.; Rodriguez, E.; Ashley, G.; Viswanathan, N.; Carney, J.; Santi, D. V.; Hutchinson, C. R.; McDaniel, R. *Biochemistry* **2003**, *42*, 72–79.
(50) Varon, M.; Fuchs, N.; Monosov, M.; Tolchinsky, S.; Rosenberg, E. *Antimicrob. Agents Chemother.* **1992**, *36*, 2316–2321.
(51) Andersen, J. F.; Hutchinson, C. R. *J. Bacteriol.* **1992**, *174*, 725–735.
(52) Molnar, I.; Schupp, T.; Ono, M.; Zirkle, R.; Milnamow, M.; Nowak-Thompson, B.; Engel, N.; Toupet, C.; Stratmann, A.; Cyr, D. D.; Gorlach, J.; Mayo, J. M.; Hu, A.; Goff, S.; Schmid, J.; Ligon, J. M. *Chem. Biol.* **2000**, *7*, 97–109.
(53) Chen, H.; Walsh, C. T. *Chem. Biol.* **2001**, *8*, 301–312.
(54) Chen, H.; Hubbard, B. K.; O'Connor, S. E.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 103–112.
(55) Chen, H.; Thomas, M. G.; O'Connor, S. E.; Hubbard, B. K.; Burkart, M. D.; Walsh, C. T. *Biochemistry* **2001**, *40*, 11651–11659.
(56) Schwarzer, D.; Finking, R.; Marahiel, M. A. *Nat. Prod. Rep.* **2003**, *20*, 275–287.
(57) Schneider, T. L.; Shen, B.; Walsh, C. T. *Biochemistry* **2003**, *42*, 9722–9730.
(58) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
(59) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
(60) Hanson, R. L.; Singh, J.; Kissick, T. P.; Patel, R. N.; Szarka, L. J.; Mueller, R. H. *Bioorg. Chem.* **1990**, *18*, 116–130.

NP0704250