J. Biochem. 127, 779–789 (2000)

Polyketide Synthase Gene Coupled to the Peptide Synthetase Module Involved in the Biosynthesis of the Cyclic Heptapeptide Microcystin[1]

Tomoyasu Nishizawa,* Akiko Ueda,* Munehiko Asayama,* Kiyonaga Fujii,† Ken-ichi Harada,† Kozo Ochi,‡ and Makoto Shirai*[,2]

*Division of Biotechnology, School of Agriculture, Ibaraki University, Ami, Ibaraki 300-0393, †Faculty of Pharmacy, Meijo University, Tempaku, Nagoya 468-8503; ‡National Food Research Institute, Tsukuba, Ibaraki 305-8642, and §Gene Research Center, Ibaraki University, Ami, Ibaraki 300-0393

Received January 17, 2000, accepted February 11, 2000

The peptide synthetase gene operon, which consists of *mcyA*, *mcyB*, and *mcyC*, for the activation and incorporation of the five amino acid constituents of microcystin has been identified [T. Nishizawa *et al.* (1999) *J. Biochem.* 126, 520–529]. By sequencing an additional 34 kb of DNA from microcystin-producing *Microcystis aeruginosa* K-139, we identified the residual microcystin synthetase gene operon, which consists of *mcyD*, *mcyE*, *mcyF*, and *mcyG*, in the opposite orientation to the *mcyABC* operon. McyD consisted of two polyketide synthase modules, and McyE contained a polyketide synthase module at the N-terminus and a peptide synthetase module at the C-terminus. McyF was found to exhibit similarity to amino acid racemase. McyG consisted of a peptide synthetase module at the N-terminus and a polyketide synthase at the C-terminus. The microcystin synthetase gene cluster was conserved in another microcystin-producing strain, *Microcystis* sp. S-70, which produces Microcystin-LR, -RR, and -YR. Insertional mutagenesis of *mcyA*, *mcyD*, or *mcyE* in *Microcystis* sp. S-70 abolished microcystin production. In conclusion, the *mcyDEFG* operon is presumed to be responsible for 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda) biosynthesis, and the incorporation of Adda and glutamic acid into the microcystin molecule.

Key words: cyanobacteria, microcystin biosynthesis, multifunctional enzyme complex, peptide synthetase gene, polyketide synthase gene.

The mass production of cyanobacteria (blue-green algae), oxygenic phototrophs, occurs world wide in eutrophic water bodies. A serious problem is the production of potent cyclic hepatotoxins, termed microcystins, in waterblooms (*1*, *2*). Microcystins are produced by several cyanobacterial genera, *Microcystis*, *Anabaena*, *Oscillatoria*, and *Nostoc*, with *Microcystis* spp. are probably the most deleterious freshwater bloom-forming cyanobacteria. In 1996, liver failure and death after exposure to microcystin were reported (*3*). Microcystins, which are potent inhibitors of protein phosphatases 1 and 2A, cause cytokeratin hyperphosphorylation, which leads to the disruption of cytoskeletal components and to cell deformation, followed by disruption of the liver architecture (*1*, *4*). Moreover, microcystins have been reported to hasten tumor development (*2*, *5*).

The general structure of microcystins is cyclo (–D-Ala–X–D-MeAsp–Z–Adda–D-Glu–Mdha–), in which X and Z are various L-amino acids, Adda is 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid, D-MeAsp is D-*erythro*-β-*iso*-asparic acid, and Mdha is *N*-methyl-dehydroalanine (*2*) (Fig. 1). Over 50 structural variations of microcystins have been isolated from cyanobacteria. Peptides containing non-protein amino acids are synthesized non-ribosomally by a large multifunctional enzyme complex, utilizing a thio-template mechanism, called non-ribosomal peptide synthetase (NRPS) (*6–8*). In 1997, Dittmann *et al.* cloned and identified a part of the microcystin synthetase genes from *Microcystis aeruginosa* PCC7806 (*9*). Recently, we identified the microcystin synthetase gene operon (*mcyA*, *B*, and *C*) including five amino acid activation modules from *M. aeruginosa* K-139, which produced 7-desmethyl-microcystin (MCYST)-LR and 3,7-didesmethyl-MCYST-LR (*10*). A gene disruption experiment revealed that the *mcy* gene is responsible for the microcystin production by K-139 cells. The arrangement of amino acid activation modules suggests that the cloned genes are responsible for activation of Mdha, D-Ala, L-Leu, D-MeAsp, and L-Arg. However, the complete gene structure for microcystin biosynthesis has yet to be identified.

Adda in the microcystin molecule is an amino acid with an unusual and modified structure. Moore *et al.* reported that Adda is biosynthesized from the carbon skeleton of phenylalanine and four molecules of acetate (*11*). These results suggest that the synthesis of Adda is catalyzed by a

[1] This work was supported by Grants for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan.
[2] To whom correspondence should be addressed. Tel +81-298-88-8652, Fax· +81-298-88-8653, E-mail· shirai@ipc.ibaraki.ac.jp
Abbreviations. MCYST, microcystin, D-MeAsp, D-*erythro*-β-methyl-aspartic acid, Adda, 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid; Mdha, *N*-methyldehydroalanine; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase, KS, β-ketoacyl-ACP-synthase; AT, acyltransferase; ACP, acyl carrier protein; KR, β-ketoacyl-ACP-reductase, DH, dehydratase; ER, enoyl reductase; SAM, *S*-adenosylmethionine; GSA, glutamate-1-semialdehyde aminotransferase, Rac, racemase; Cm, chloramphenicol.

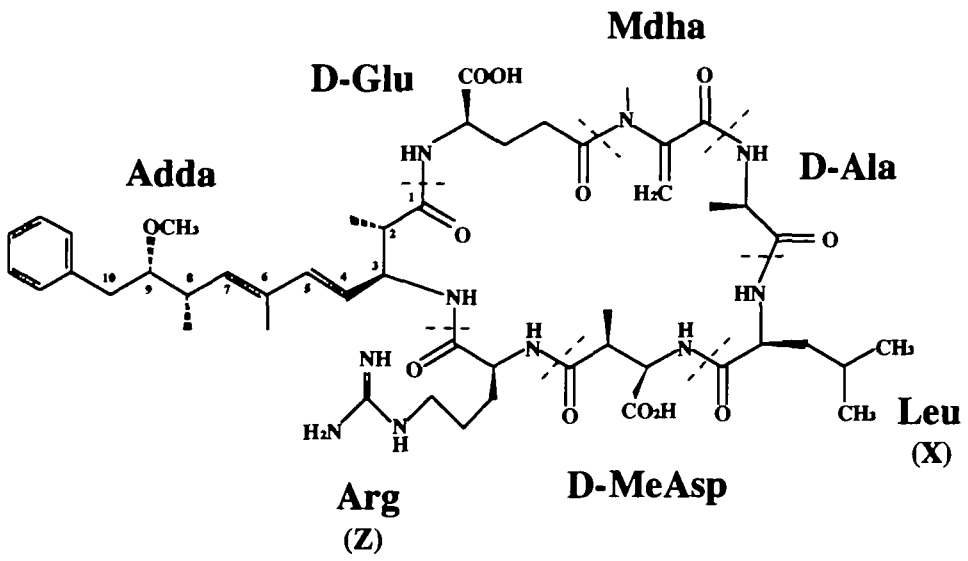

Fig 1 **General structure of microcystin.** The general structure of microcystin (MCYST-XZ) is cyclo (–D-Ala–X–D-MeAsp–Z-Adda-D-Glu–Mdha–), in which X and Z are various L amino acids, D-MeAsp is D-*erythro*-β-methylaspartic acid, Adda is 3-amino-9-methoxy-10-phenyl-2,6,8-trimethyl-deca-4,6-dienoic acid, and Mdha is *N*-methyldehydroalanine This figure shows microcystin-LR (MCYST-LR)

modular polyketide synthase multi-enzyme complex. Furthermore, we detected the open reading frame, which showed high degrees of similarity to modular polyketide synthases, upstream of *mcyA* (*10*). These findings encouraged us to investigate the polyketide synthase (PKS) gene in microcystin-producing cells.

Polyketide formation is analogous to fatty acid synthesis (FAS) (*12–14*). PKS systems are classified into two types (*15*, *16*): type I PKSs (modular PKSs) are large multifunctional enzymes with a unique modular structure in which each module is responsible for the activation, initiation, elongation and termination steps. Type II PKSs are systems made up of individual enzymes.

In this study, we found a unique PKS module coupled to a peptide synthetase module and revealed that the PKS genes were responsible for microcystin production.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Culture Conditions, and Plasmids—All strains and plasmids used in this work are described in Table I. *Microcystis* strains were grown under 2,000-lx continuous illumination from fluorescent (cool white) lighting at 30°C in CB medium (*17*). *Escherichia coli* DH5αMCR (Cosmo Bio., Tokyo) was used as a host for recombinant plasmids and grown at 37°C for 16 h in LB broth. When necessary, antibiotics were added at the following final concentrations: 75 μg/ml ampicillin and 30 μg/ml neomycin. Lorist6 DNA (Nippon Gene, Toyama), for the cosmid library, and pBluescript II KS+ and pUC118/119 (TOYOBO), for cloning, were used.

DNA Manipulation—Total DNA of *Microcystis* strains was isolated from cells grown to a late logarithmic phase by means of the previously described procedure (*18*). DNA manipulations were performed as described (*19*).

Cloning of the mcy Genes—Total DNA of *M. aeruginosa* K-139 was digested with *Cla*I, *Xba*I, *Hin*cII, or *Hin*dIII and then inserted into the pBluescript II KS+ phagemid vector or the pUC vector, generating plasmid-libraries. At first, using a 1.5 kb *Eco*RI fragment from the 3′-end of the insert of pCOTnβ5 as a probe, a recombinant pKCA2 containing a 3.5 kb *Cla*I fragment was isolated (Fig. 2). Next, a cosmid pCOTnβ13 containing a 14.4 kb fragment was isolated from a Lorist6 DNA genomic library (*10*) using a 1.6 kb *Cla*I–*Eco*RI fragment from the 3′-end of the insert of pKCA2 as a probe. The following recombinant plasmids were isolated from the plasmid libraries using appropriate DNA fragments as probes (the probes are indicated in parentheses); pKCB1 containing a 5.0 kb *Cla*I fragment (a 0.5 kb *Cla*I–*Xba*I fragment from pCOTnβ13), pMCQ3 containing a 6.0 kb *Xba*I fragment and pMCP1 containing a 9.1 kb *Hin*dIII fragment (a 0.3 kb *Eco*RI fragment from pKCB1), pMCN5 containing a 2.3 kb *Xba*I fragment (a 1.8 kb *Hin*dIII–*Xba*I fragment from pMCP1), pMCM1 containing a 1.9 kb *Hin*cII fragment (a 1.0 kb *Hin*cII–*Xba*I fragment from pMCN5), pMCL9 containing a 2.4 kb *Hin*dIII fragment (a 1.0 kb *Hin*cII–*Xba*I fragment from pMCM1), and pMCK3 containing a 4.3 kb *Xba*I fragment (a 1.3 kb *Xba*I–*Hin*dIII fragment from pMCL9).

Construction of Gene Disruption Plasmids and Integrative Conjugation of Microcystis—Plasmids for gene disruption of *mcyB*, *mcyD*, and *mcyE* by homologous recombination were constructed as follows. The 4,325 bp *Xba*I–*Hin*dIII fragment containing an amino acid activation domain of *mcyB* from pCOTnβ5 was cloned into the *Xba*I–*Hin*dIII sites of pUC119, generating pMCWX. The 1.2 kb *Sma*I fragment containing the *Cm*^{r} gene cassette from pR107XH and the 1.8 kb *Bam*HI fragment containing the *mob* gene from pSUP5011 were inserted into the *Eco*RV site (in *mcyB*) and *Hin*dIII site (at the multicloning site) of pMCWX, respectively, generating pFXS3. The 1,699 bp *Xba*I–*Hin*dIII fragment containing the DH-KR domain of module 6/*mcyD* from pKCA2 was inserted into the *Xba*I–*Hin*dIII sites of pUC118, generating pNOR-f. A *Hin*cII fragment of the *Cm*^{r} gene cassette and the *mob* fragment were inserted into the *Hin*cII site (in *mcyD*) and *Bam*HI site of pNOR-f, respectively, generating pEXS5. The 3,636 bp *Bgl*II fragment containing an amino acid activation domain of module 10/*mcyE* from pKCB1 was cloned into the *Bam*HI site of the pNC122 vector, generating pNCB54. The *Xba*I fragment of the *Cm*^{r} gene cassette and the *mob* fragment were inserted into the *Xba*I site (in *mcyE*) and *Bgl*II site (at the

TABLE I **Strains and plasmids.**

Strain/Cosmid/Plasmid	Relevant characteristic	Reference
Strain		
M. aeruginosa		
B-19	wild type, non-microcystin producing	(*18*)
K-81	wild type, non-microcystin producing	(*30*)
K-139	wild type, microcystin producing	(*42*)
S-77	wild type, microcystin producing	(*17*)
Microcystis sp.		
S-70	wild type, microcystin producing	(*18*)
DX4, 5, and 6	S-70, but *mcyA*::Cm by conjugation with pDXS7	This study
EX1	S-70, but *mcyD*::Cm by conjugation with pEXS5	This study
FX1 and 2	S-70, but *mcyB*::Cm by conjugation with pFXS3	This study
GX1 and 2	S-70, but *mcyE*::Cm by conjugation with pGXS3	This study
Escherichia coli		
DH5αMCR	*λ-F-supE44 hsdR17 Δ(lacZAY-argF)U169(ϕ80lacZΔM15) recA1 mcrAΔ(mrr hsdRMS mcrBC) deoR*	Cosmo Bio
S17-1	*Δres, mod*$^+$, *thi, pro, hsdR, recA*, RP4 derivative integrated, Spe^r	(*43*)
Cosmid		
Lorist6 DNA	Cloning vector, Neo^r, cloning sites *Hin*dIII, *Not*I, *Sca*I, *Bam*HI	Nippon Gene
pCOTnβ library	c. 40-kb *Hin*dIII fragment from *M. aeruginosa* K-139 genomic DNA, on Lorist6 DNA	(*10*)
Plasmid		
pBluescript II KS+	Cloning vector, Amp^r	TOYOBO
pUC119	Cloning vector, Amp^r	TOYOBO
pNC122	Cloning vector, Amp^r, pUC119 containing a *Bgl*II site between *Pst*I and *Sph*I sites, and a *Nco*I site between *Pst*I and *Bam*HI, but removing *Hin*cII and *Xba*I sites in the multiple cloning sites	This study
pKCA2	6.2-kb *Cla*I fragment from *M. aeruginosa* K-139, on pBluescript II KS+ (refer to Fig. 2)	This study
pKCB1	5.0-kb *Cla*I fragment from K-139, on pBluescript II KS+ (refer to Fig. 2)	This study
pMCK3	4.3-kb *Xba*I fragment from K-139, on pUC119 (refer to Fig. 2)	This study
pMCL9	2.4-kb *Hin*dIII fragment from K-139, on pUC119 (refer to Fig. 2)	This study
pMCM1	1.9-kb *Xba*I fragment from K-139, on pUC119 (refer to Fig. 2)	This study
pMCN5	2.3-kb *Hin*cII fragment from K-139, on pUC119 (refer to Fig. 2)	This study
pMCP1	9.1-kb *Hin*dIII fragment from K-139, on pUC119 (refer to Fig. 2)	This study
pMCQ3	6.0-kb *Xba*I fragment from K-139, on pUC119 (refer to Fig. 2)	This study
pSUP5011	Mobilizable plasmid derived from pBR325	(*43*)
pR107XH	1.2-kb Cm^r gene cassette with the *Microcystis rpoD1* promoters, on pUC118	(*10*)
pDXS7	*mcyA* gene disruption on plasmid, 2.7 kb *Hin*cII fragment with Cm^r cassette and mob, on pUC119	(*10*)
pEXS5	*mcyD* gene disruption on plasmid, 1.7 kb *Hin*dIII and *Xba*I fragment with Cm^r cassette and mob, on pUC118	This study
pFXS3	*mcyB* gene disruption on plasmid, 4.3 kb *Hin*dIII and *Xba*I fragment with Cm^r cassette and mob, on pUC119	This study
pGXS3	*mcyE* gene disruption on plasmid, 3.6 kb *Bgl*II fragment with Cm^r cassette and mob, on pNC122	This study

multicloning site) of pNCB54, respectively, generating pGXS13. The plasmids for gene disruption were introduced into *Microcystis* cells by conjugation from *E. coli* S17-1 and chloramphenicol resistant (8 μg of chloramphenicol per ml) conjugants were selected, as described previously (*10*).

Southern Hybridizations—Digested cyanobacterial DNA was separated on 0.8% agarose gels and then transferred to Amersham Hybond-N or NX membranes as described previously (*10*). DNA fragments, as probes, were labeled using an ECL random prime labeling kit (Amersham Pharmacia Biotech). Southern hybridization and detection were performed as recommended by the manufacturer.

HPLC Analysis of Microcystins—Microcystins were extracted from dried cells with 5% aqueous acetic acid, cleaned up using a Bond Elute ODS cartridge (Varian, CA), and then analyzed by HPLC as described previously (*10*).

DNA Sequencing and Computer Analysis—The nucleotide sequence was determined by dideoxy chain termination, using an Applied Biosystems Automated Sequencer (model 373S) (*10*). The DNA sequences were assembled and analyzed using GENETYX-MAC software from Software Development (Tokyo).

Nucleotide Sequence Accession Number—The nucleotide sequences in this report have been submitted to DDBJ under the following accession number: AB032549 (*mcyD, mcyE, mcyF,* and *mcyG* from *M. aeruginosa* K-139).

RESULTS

Cloning of mcyD, mcyE, mcyF, and mcyG—The *mcyA, B,* and *C* genes encoding five modules that activate the five amino acid constituents, and a putative open reading frame 743 bp upstream of *mcyA*, which was in the opposite orientation to *mcyABC*, were found (Fig. 2) (*10*). This open reading frame shows high similarity to PKS of *Mycobacterium* (*20*). To obtain the microcystin synthetase genes, a 34 kb region including PKS genes was cloned from *M. aeruginosa* K-139 and sequenced on both strands, as described under "EXPERIMENTAL PROCEDURES." The sequence had a typical GC-content of about 40%, this value being almost the same as that for *M. aeruginosa* K-139 genomic DNA (*18*).

Analysis of the cloned nucleotide sequence revealed three huge open reading frames (ORFs) and three small ORFs with the same direction of transcription (Fig. 2). The first ORF, *mcyD*, is 11,718 bp in length, encoding a polypeptide of 3,906 amino acids (aa) with a predicted molecular mass of 435,915 Da. The putative Shine-Dalgarno sequence (AAG–GA) was found 9 nucleotides upstream of the start codon. The second ORF, *mcyE*, is located 167 bp downstream of the TAA stop codon of *mcyD*. A possible Shine-Dalgarno sequence (AGAGAA) is located 6 bp upstream of

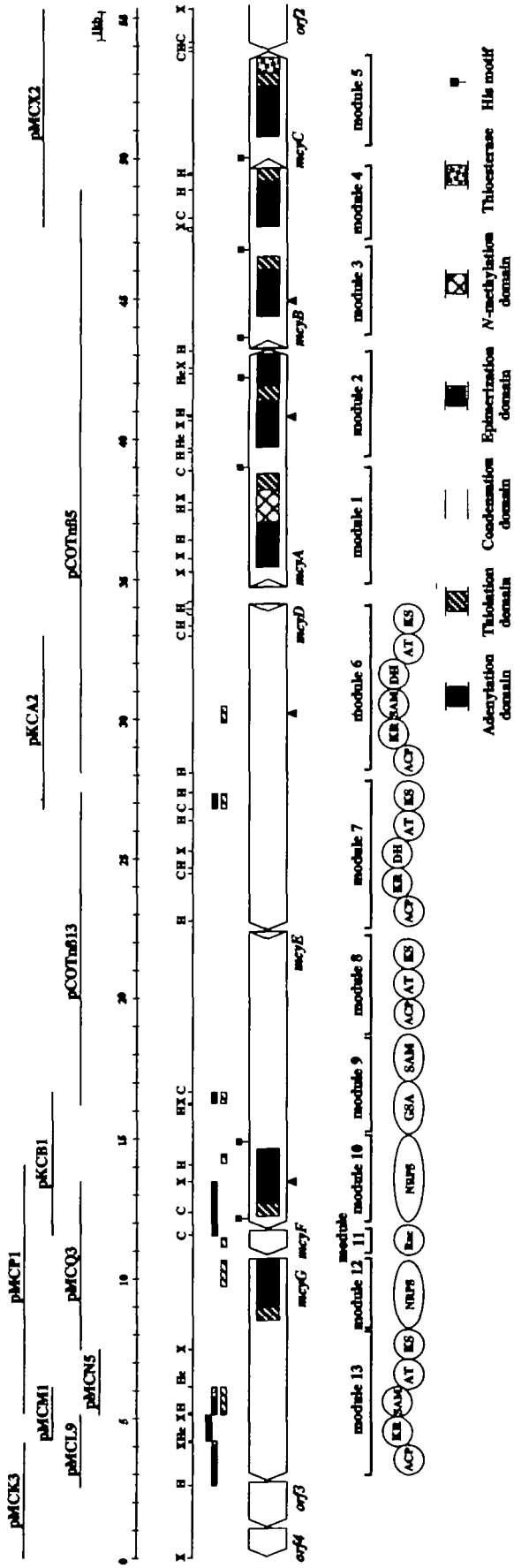

Fig 2. **Organization and physical map of the *mcy* gene cluster.** Modules of the microcystin synthetase are shown The structure of a typical module including the domains and His motif conserved in peptide synthetase are illustrated in the box. Each circle represents an enzymatic domain in the polyketide synthase multifunctional protein. KS, β-ketoacyl-ACP synthase; ACP, acyl carrier protein, AT, acyltransferase, KR, β-ketoacyl-ACP reductase; DH, dehydratase, SAM, *S*-adenosylmethionine–dependent methyltransferase GSA and Rac represent glutamate-1-semialdehyde aminotransferase and racemase, respectively NRPS is non-ribosomal peptide synthetase. The restriction sites and numbers of modules are shown above and below the diagram, respectively The upper part of the figure shows the locations of cosmids and plasmids. The striped and stippled fields above the diagram shows the positions corresponding to the probes used for cloning and Southern hybridization, respectively. The closed triangles show the insertion sites of the Cm cassette for gene disruption Abbreviations: C, *Cla*I; Hc, *Hin*cII; H, *Hin*dIII; X, *Xba*I

the ATG codon. This ORF (10,461 bp) encodes a putative protein of 3,487 aa with a predicted molecular mass of 392,319 Da. The third ORF, *mcyF*, is located 35 bp downstream of *mcyE* and is 753 bp in length, encoding a 251 residue polypeptide with a predicted molecular mass of 27,990 Da. The putative Shine-Dalgarno sequence (AGGAGA) was found 4 nucleotides before the putative initiation codon. The fourth ORF, *mcyG*, is located 74 bp downstream of *mcyF* and is 7,896 bp in length, encoding a 2,632 residue polypeptide with a predicted molecular mass of 294,299 Da. The putative Shine-Dalgarno sequence (AAGAGG) was found 10 nucleotides before the putative initiation codon. Moreover, we observed another putative ORF (*orf3*), 85 bp downstream of *mcyG* and following *orf4*, which were in the same orientation.

Homology Analysis of the Cloned Genes—Four modules for type I PKS (modules 6, 7, 8, and 13) were identified on the cloned genes on comparison with the PKS biosynthesis domains of the 6-deoxyerythronolide B synthase (DEBS) of *Saccharoplyspora erythrase* (*12*). Interestingly, a PKS module coupled to the NRPS module was found (Fig. 2). PKSs have a unique modular structure in which each module is responsible for the catalysis of one cycle of polyketide chain elongation, which resembles the steps in FAS (*12*, *21*). Each module can be subdivided into specific domains, β-ketoacyl-ACP-synthase (KS), acyltransferase (AT), and acyl carrier protein (ACP). In addition, β-ketoacyl-ACP-reductase (KR), dehydratase (DH), and enoyl reductase (ER) domains are present. McyD consists of two PKS modules, module 6 including an extent of KS, AT, DH, SAM, KR, and ACP, and module 7 including an extent of KS, AT, DH, KR, and ACP (Fig 2). McyE contains module 8 including an extent of KS, AT and ACP, and the fourth ORF, McyG, contains module 13 including an extent of KS, AT, SAM, KR, and ACP. The minimal domains of the four modules in *Microcystis* PKS show similarity (37.0–54.1% in the KS domain, 25.9–37.5% in the AT domain, and 15.7–35.4% in the ACP domain).

Generally, conserved amino acid motifs are observed in each domain of PKS modules (*12*, *21*, *22*). Alignment of the amino acid sequences of conserved motifs is shown in Fig. 3. In KS domains of Mcy, active site motifs and the catalytic cysteine residue in the presumed active-site domain were conserved (Fig. 3A). Alignment of the substrate specific sequence motifs and the consensus sequence of the active site in the ATs (*23*) is shown in Fig. 3B. The signature sequence [GHSxG] was conserved in the four ATs of Mcy.

However, it is difficult to infer the substrate, *i.e.* acetate or propionate, from the sequences. The ACP domain of type I and type II PKSs contains the consensus signature [LGx-DS] of the presumed binding site for the cofactor 4′-phosphopantetheine (*24*). The active-site serine and surrounding amino acid sequence were conserved among four Mcys (Fig. 3C). DH and KR, or KR domains are present between the AT and ACP domains of modules 6, 7 and 13 (Fig. 2), and show similarity (29.3% in the DH domain, and 33.3–41.0% in the KR domain). Alignment of the sequences showed that the DH domains in Mcy have the apparent active site motif [HxxxD/ExxxxP] and that the KR domains contain a potential motif [TGGxGxxGxxxA/T] for NADP(H) binding (Fig. 3D).

The amino acid sequences of the internal regions between the DH and KR domains of module 6, and the AT and KR domains of module 13, and the downstream region of the ACP domain in module 8 showed local similarity to a part of the N-terminus of *N*-methyltransferases (Fig. 4A). The glycine-rich sequence [VL(E/D)xGxGxG] has been suggested to be a possible component of a binding site for *S*-adenosylmethionine (SAM) (*25*). The sequence identity of these SAM domains (about 460 aa) in modules 6, 9, and 13 is 29.5–32.7%. A computer homology search of module 9 of *mcyE* with available data bases revealed that a downstream region of SAM showed significant similarity to glutamate-1-semialdehyde aminotransferase (GSA). The GSA domain of McyE, which is composed of about 400 aa, shows about 30% identity to the GSAs from barley, *Nicotiana tabacum*, soybean, *E. coli*, and *Synechococcus* (*26*). A putative pyridoxamine phosphate binding lysine and surrounding amino acid sequences are conserved in module 9 of McyE (Fig. 4B).

Interestingly, an NRPS module was observed at the C-terminus of McyE and at the N-terminus of McyG (Fig. 2). These two modules consist of adenylation and thioester-binding domains, and show a high degree of amino acid sequence identity (30.4–35.9% in the adenylation domain, and 29.1–45.4% in the thiolation domain). Sequence alignment of the conserved core motifs of Mcy, core1–core6 (*7*, *27*, *28*), is shown in Fig. 4C. The consensus core sequences are highly conserved in modules 10 and 12, except for core1 of module 12.

McyF shows 30.4% identity to the aspartate racemase of the sulfur-dependent hyperthermophilic archaeum *Desulfurococcus* Strain SY, which is quite widespread in eubacteria as the D-amino acid constitutes parts of the fundamental tetrapeptide chain in murein of the cell wall (*29*). Cysteine residues at the active-site, which are thought to be the catalytic center of these cofactor independent racemases, were conserved in motifs 1 and 2 of McyF (Fig. 4D).

Moreover, a computer homology search showed that Orf3 exhibits similarity to the hypothetical ABC transporter ATP-binding protein (sll0182) of *Synechosystis* PCC6803 and Orf4 to the D-3-phosphoglycerate dehydrogenase of *Methanobacterium thermoautotrophicum*, respectively.

Presence of the mcy Genes in Other Microcystis Strains—The presence of *mcyD*, *mcyE*, *mcyF*, and *mcyG* in other *Microcystis* strains was examined by genomic Southern hybridization. The chromosomal DNAs from *Microcystis* strains were digested with *Cla*I, *Hin*dIII, or *Xba*I, and then Southern hybridization was performed using the module-specific probes shown in Fig. 2. All signals were detected in the microcystin-producing strains *M. aeruginosa* K-139 and S-77, and *Microcystis* sp. S-70, but not in the microcystin-non-producing strain *M. aeruginosa* B-19 (Table II). Interestingly, non-toxic *M. aeruginosa* K-81 strain (LD_{50}, >1,000 mg of dry cells per kg of mouse-weight) (*30*) was positive for all probes.

Disruption of the mcyD and mcyE Genes—To obtain conclusive proof of the involvement of *mcyD* and *mcyE* in microcystin production, insertional mutagenesis of *mcyD*

A)

module	active site motif of KS domain	
McyD/mod.6	GPSLAVDTACSSSLVAVH	
McyD/mod.7	GPSMTIDTACSSSLVAIH	
McyE/mod.8	GPCLSIDAACASSLAAVH	
McyG/mod.13	GPSVHVQTACSTGLVVVH	
DEBS/mod.1	GPAISVDTACSSSLVAVH	PKS type I
RAPS/mod.1	GPAITVDTACSSSLVALH	PKS type I
RifA/mod.1	GPAVTVDTACSSSLCAMH	PKS type I
Act I/orf1	GPVTMVSTGCTSGLDSVG	PKS type II
Tcm I/orf1	GPVTVVSTGCTSGLDAVG	PKS type II

B)

module	AT domain motif			active site
McyD/mod.6	ISLSVQPVLFAYQYALCELW	- 11aa -		GSGLG
McyD/mod.7	ETQITQPVIFSLEYALAKLW	- 11aa -		GHSIG
McyE/mod.8	QTAYAQPAIFALEYSLTHLW	- 11aa -		GHSVG
McyG/mod.13	QTDYTQPALFLIEVALAQLW	- 11aa -		GHSLG
Propionate ATs	RVDVVxxxxxxxMxSxAxhW	- 11aa -		GHSxG
Acetate ATs	ETGYAxxxxxxxQxAxFGLL	- 11aa -		GHSxG

C)

module	active site motif of ACP domain	
McyD/mod.6	FFDLGMDSLTSTE	
McyD/mod.7	FLDLGLNSLMVIE	
McyE/mod.8	LLNLGADSIILTD	
McyG/mod.13	LIEMGIDSLSSIE	
DEBS/mod.1	FAELGVDSLSALE	PKS type I
RAPS/ mod.1	FRDLGVDSLTAVE	PKS type I
RAPS/mod.11	FKDLGINSLTAVE	PKS type I
RifA/mod.1	FKDAGFDSLTAVE	PKS type I
Act I/orf3	FEDIGYDSLALME	PKS type II
Tcm I/orf3	YQDLGYDSIALLE	PKS type II

D)

module	active site motif	
	DH	KR
McyD/mod.6	HQVFDQAILP	TGGTGGLGLATT
McyD/mod.7	HKVFENIVFP	TGGGGKLGGLVA
McyG/mod.13	----------	TGGIGHLGLELA
DEBS/mod.1	HVVQGRTCVP	TGGTGGVGGQIA
RAPS/mod.1	HAVRGSVLLP	AGGVGMAATQIA
RifB/mod.4	HAIGGVVLIP	TGGTGSLGGLVA
consensus sequence	HxxxGxxxxP	TGGxGxxGxxxA

Fig. 3. **Alignment of conserved motifs among polyketide synthases.** (A) Comparison of KS domains. (B) Comparison of AT domains Sequence motifs for the acetate and propionate ATs are shown under the sequence alignment. (C) Comparison of ACP domains (D) Comparison of DH and KR domains. The consensus sequences for DH and KR domains are shown. DEBS, deoxyerythronolide B synthase (*12*); RAPS, rapamycin synthase (*21*); RifA and RifB, rifamycin B biosynthesis (*22*), Act I, actinorhodin synthase (*44*); Tcm I, tetracenomycin C synthase (*45*)

A)

enzyme/module	sequence
Esyn	VLEIGTGSGMILFN
CYSYN/mod.2	VLEIGTGTGMVLFN
McyA/mod.1	VLEIGCGTGLILFQ
McyE/mod.6	ILEIGGGTGAATTG
McyE/mod.9	ILEIGGGTGATSEA
McyG/mod.13	ILEIGAGTGATTQQ
EryG	VLDVGFGLGAQDFF

B)

strain	sequence
barley	TTLGKIIGGGLP
N. tabacum	TTLGKIIGGGLP
soybean	TTLGKIIGGGLP
E. coli	TCLGKIIGGGMP
Synechococcus	TTLGKIIGGGLP
McyE/mod.9	VIYGKAIGGGLP

C)

function	core 1 unknown	core 2 ATP binding	core 3 ATPase motif	core 4 ATP binding	core 5 ATP binding	core 6 4'-phosphopantetheine binding
[motif]	LKAGxAYVPLD L I	LAYxxYTSGSTGxPKG T	GELxIxGxGVARGYL L	YRTGDL K	GRxDxQVKIRGxRIELGEIE	DxFFxLGGHSL D I
McyA/mod.1	LKAGGAYVPLD	LAYVIYTSGSTGKPKG	GELHIGGIGLARGYL	YKTGDL	GRIDHQVKIRGFRIELGEIE	DNFFELGGDSI
McyA/mod.2	LKAGGAYVPLD	TAYVIYTSGSTGKPKG	GDLYIGGKGLARCYL	YKTGDL	GRGDNQVKLRGFRIELGEIE	DRFFEVGGDSI
McyB/mod.3	LKAGGAYVPLD	LAYVIYTSGSTGKPKG	GELHISSVGLARGYL	YKTGDL	GRIDNQVKLRGLRIELGEIE	DNFFELGGHSL
McyB/mod.4	FKAGGIYLPLR	SNYIMFTSGSTGEPKA	GEVFVKSPYLTKGYY	YRTGDL	GRSDNQIKLNGVRIELGEIE	QSFFELGGNSL
McyC/mod.5	IKAGGAYVPLD	LAYVIYTSGSTGKPKG	GELHIGGMGLARGYL	YKTGDL	GRIDNQVKLRGLRIELGEIQ	DNFFELGGHSL
McyE/mod.10	FKTGATYVPID	TAY-IFTSGSTGNPKG	GEICVEGAALASGYH	FRTGDL	GRKDNQVKVNGYRIDPGEIE	DNFFEIGGHSL
McyG/mod.12	FLGGIIPVPLT	AALLLFTSGSTGMPKG	GRLQIQGHSVTKGYY	FTTGDL	GREKQEIIINGVNYFAHELE	DNFFELGGHSI
[concensus]	-kaGgayvPld	-ayviyTSGSTGkPKg	Gel-i-g-glargYl	ykTGDL	GR-dnqvElrG-rielgEIe	dnFFElGG-Sl

D)

type	strain	active site motif 1	active site motif 2
Asp	*Desulfrococcus*	DFIIMPCNTAHFFAED	GIIAGCTEVSVAL
	Streptococcus thermophilus	NFIVLTCNTAHYFFEE	KVILGCTELSLMN
	Pyrococcus horikoshii	DFIIMPCNTAHAFVED	CIIAGCTEVSVVL
Glu	*Lactobacillus brevis*	KALVIACNTATAAALT	TLVMGCTHFPLIR
	Bacillus subtilis	KMLVIACNTATAIALD	TVILGCTHYPILQ
	Escherichia coli	ALAVVACNTASTVSLP	TVVLGCTHFPLLQ
	Microcystis aeruginosa K-139	DCIVIGCCTAHYALPQ	SYISGCTEFHLLT

Fig 4 **Alignment of conserved motifs in the functional domains.** (A) Common regions in *S*-adenosylmethionine–dependent methyltransferases. The conserved sequence motifs are shown in bold Esyn, enniatin synthetase (*25*); CYSYN, cyclosporin A synthetase (*46*), McyA, microcystin synthetases (*10*), EryG, erythromycin synthase (*47*) (B) The conserved amino acid sequences surrounding the putative pyridoxal binding lysine (bold letter) among the six glutamate 1-semialdehyde aminotransferases (GSA) Barley, M31545, *N tabacum*, X65974; soybean, L12453; *E. coli*, X53696; *Synechococcus* PCC6301, X53695, McyE/mod.9, AB032549 (C) The alignment of core sequences of peptide synthetase modules in microcystin synthetase. Putative function and consensus motif sequences are shown at the top, and consensus sequences in Mcy modules at the bottom McyA, McyB, and McyC, microcystin synthetases (*10*). (D) The alignment of acidity amino acid racemases. The presumed catalytic and conserved residues are indicated in bold. *Desulfurococcus*, D84067; *Streptococcus thermophilous*, X61301; *Pyrococcus horikoshii*, AP000003; *Microcystis aeruginosa* K-139, AB032549; *Lactobacillus brevis*, D29627; *Bacillus subtilis*, Z99118; *Escherichia coli*, AE000471

TABLE II **Summary of the results of genomic Southern hybridization of Microcystis DNA with a DNA fragment from each module.**

Strain	*mcyD*		*mcyE*		*mcyF*	*mcyG*		Microcystin
	module 6	module 7	module 9	module 10	module 11	module 12	module 13	
M. aeruginosa								
B-19	−	−	−	−	−	−	−	non toxic
K-81	+	+	+	+	+	+	+	non toxic
K-139	+	+	+	+	+	+	+	[Dha^7]MCYST-LR, [D-Asp^3, Dha^7]MCYST-LR
S-77	+	+	+	+	+	+	+	MCYST-LR, RR, YR
Microcystis sp.								
S-70	+	+	+	+	+	+	+	MCYST-LR, RR, YR

−, no signal; +, signal; [Dha^7]MCYST-LR, 7-desmethylmicrocystin-LR, [D-Asp^3, Dha^7] MCYST-LR, 3,7-didesmethylmicrocystin-LR; MCYST, microcystin

and *mcyE* in the *M. aeruginosa* K-139 genome was carried out. However, no disruptants were isolated. In a previous study, we isolated gene disruption mutants of the K-139 strain with difficulty after repetitious experiments on the conjugation (*10*). The host-restriction system may interfere with homologous recombination (*31*). Genomic Southern hybridization showed the presence of *mcyD* and *mcyE* in *Microcystis* sp. S-70, which produces MCYST-LR, -RR, and -YR. Therefore, insertional mutagenesis of *mcyD* and *mcyE* in strain S-70 was performed, and then chloramphenicol-resistant conjugants, *Microcystis* sp. S-70EX1 as a disruptant of *mcyD* and S-70GX1, 2, and 3 as disruptants of *mcyE*, were isolated. To confirm the integration of the chloramphenicol (*Cm*) cassette into the target gene, genomic Southern hybridization was carried out. Total DNA from the S-70EX1 mutant was digested with *Hin*dIII and then probed with the 1.5 kb *Eco*RI fragment containing the SAM domain of module 6 [Fig. 5A (i)]. The 3.0 kb signal was detected in the wild type of *Microcystis* sp. S-70. On the other hand, only the 4.3 kb signal was detected in the EX1 mutant [Fig. 5B (i)] and the position of this signal coincided with that of a signal observed with the *Cm* probe (data not

shown). To examine the disruption of *mcyE*, total DNAs were digested with *Hin*cII and then probed with the 0.33 kb *Eco*RI fragment containing the NRPS domain of *mcyE* [Fig. 5A (ii)]. The 3.0 kb signal was detected in the S-70 wild type. While the 2.0 kb signal was detected in mutants, S-70GX1, 2 and 3 [Fig. 5B (ii)], and the positions of these signals coincided with that of a signal observed with the *Cm* probe (data not shown). To determine the roles of *mcyD* and *mcyE* in microcystin biosynthesis, we analyzed the microcystin production of these mutants by HPLC. Extracts of the *Microcystis* sp. S-70 wild strain gave three peaks for MCYST-LR, -RR, and -YR (Fig. 6A). For the EX1, and GX1, 2, and 3 mutants, all peaks for microcystins were absent (Fig. 6, B and C). These results revealed that the *mcyD* gene encoding PKSs and the *mcyE* gene encoding the PKS coupled to the NRPS are specifically involved in microcystin biosynthesis. Furthermore, our results revealed that both genes are responsible for the production of three kinds of microcystin, MCYST-LR, -RR, and -YR.

Analysis of the Microcystis sp. S-70 mcy Cluster—Genomic Southern hybridization analysis revealed that the *Microcystis* sp. S-70 strain has the *mcyA, B, C (10), D, E, F,* and *G* gene cluster. Furthermore, a PCR-amplified DNA fragment (693 bp) with the 3′-end of *mcyB* and the 5′-end of *mcyC* of *Microcystis* sp. S-70 showed strong identity (nucleotide: 98%, amino acid: 96%) to that of *M. aeruginosa* K-139 (*10*). To examine the identity of the polyketide synthase genes in strains S-70 and K-139, the *mcyD* gene was cloned from *Microcystis* sp. S-70EX1. The *Hin*dIII fragment including the *Cm* cassette from the cells was cloned into the *Hin*dIII site of pUC119 and sequenced. Sequence analysis showed that the nucleotide identity of the DH-KR region (3,078 bp) of module 6 in *mcyD* between strains K-139 and S-70 was 97%.

McyB is thought to be involved in the activation of "variable L-amino acids" located at X in the microcystin molecule shown in Fig. 1. To determine whether one or three sets of the *mcyABC* operon are responsible for the production of MCYST-LR, -RR, and -YR, gene disruption of *mcyA* and *mcyB* in the *Microcystin* sp. S-70 genome was performed by insertional mutagenesis using a disruption plasmid, pDXS7 and pFXS3, respectively. Three conjugants, S-70DX4, 5, and 6, as *mcyA* disruptants, and two conjugants, S-70FX1 and 2, as *mcyB* disruptants, were isolated and the integration of the *Cm* cassette into the target genes was confirmed by genomic Southern hybridization (data not shown). HPLC analysis revealed that the microcystins were absent in the mutants (Fig. 6D). These results revealed that the gene organization of microcystin synthetase is almost the same in strains K-139 and S-70, and that one set of the microcystin synthetase genes is responsible for the production of three kinds of microcystins, MCYST-LR, -RR, and -YR.

Phylogenetic Analysis of the Adenylation Domain—The phylogenetic relationships of the adenylation region consisting of about 70 aa upstream of core1 (LKAGGA) to about 80 aa downstream of core 5 (RIELGEIE) from gramicidin S (*32, 33*), surfactin (*34, 35*), tyrocidine (*36*), isopeni-

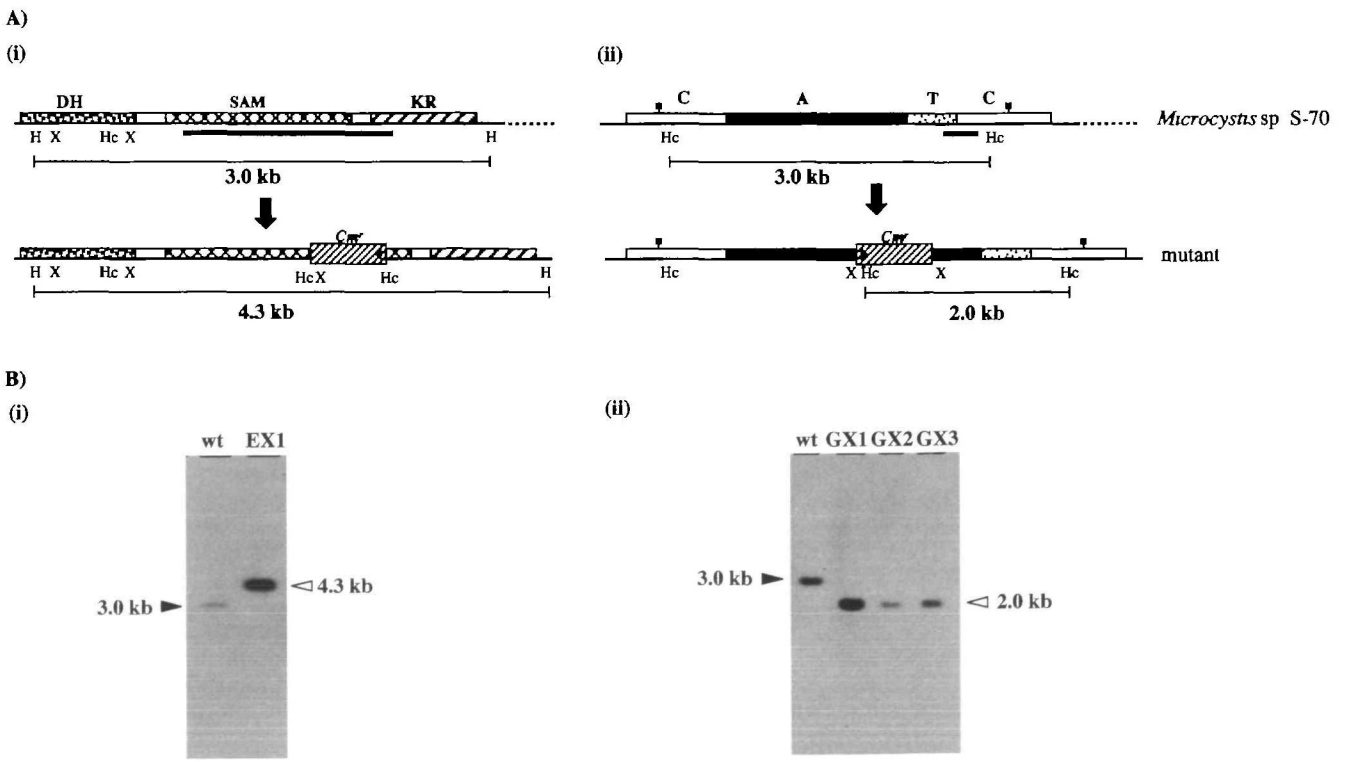

Fig. 5 **Disruption of the *mcy* genes by homologous recombination.** (A) Schematic representation of the insertional inactivation of the *mcyD* and *mcyE* genes (i) *Hin*dIII digestion fragments hybridized with a *mcyD* probe (bold bar) are shown, along with their sizes. (ii) *Hin*cII digestion fragments hybridized with a *mcyE* probe (bold bar) are shown, with their sizes. DH, dehydratase; SAM, *S*-adenosylmethionine–dependent methyltransferase; KR, β-ketoacyl-ACP reductase; A, adenylation domain, C, condensation domain; T, thiolation domain, Cm, Cm cassette, Hc, *Hin*cII, H, *Hin*dIII; X, *Xba*I (B) Genomic Southern hybridization analysis. wt, *Microcystis* sp S-70 wild type, EX1, *mcyD* mutant; GX1, 2 and 3, *mcyE* mutants. The sizes of the signals are indicated on both sides

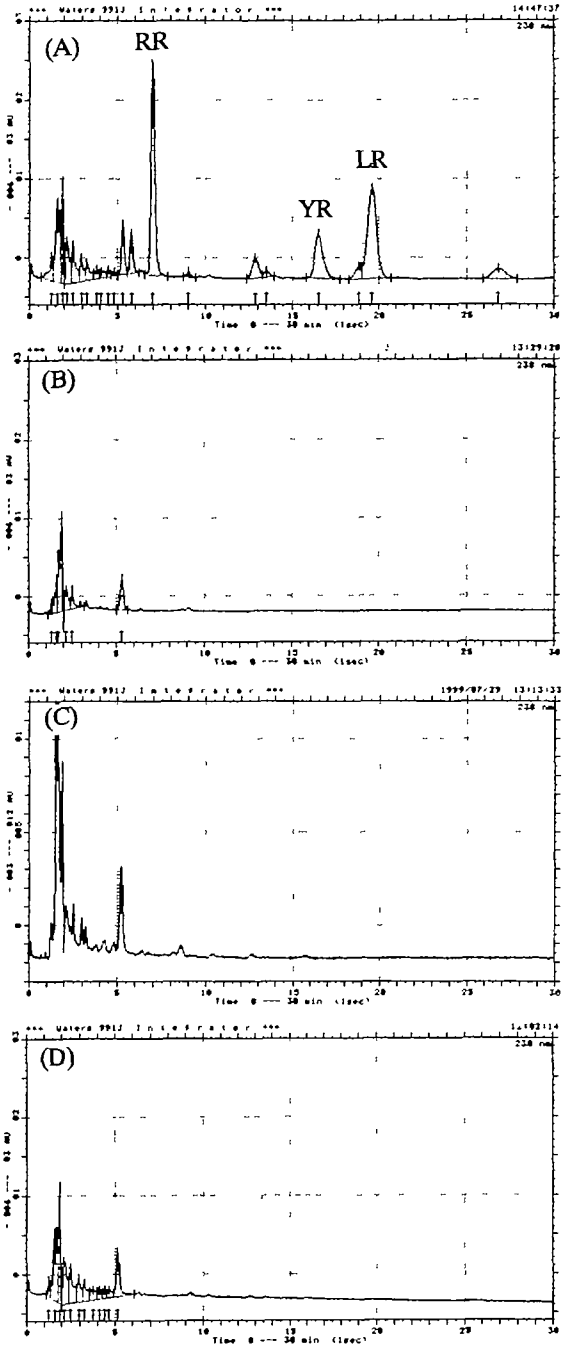

Fig. 6. **HPLC of methanol extracts of *Microcystis* sp. S-70 cells.** Microcystins were extracted from dried cells of the wild-type (A), mutant EX1 (B), mutant GX1 (C), and mutant FX1 (D) Peak RR, MCYST-RR; peak YR, MCYST-LR; peak LR, MCYST-YR; column, Cosmosil 5C18ARII (150 × 4.6mm I.D.); mobile phase, CH_3CN·0 01 M trifluoroacetic acid = 30:70 (v/v); flow rate, 1.0 ml/min; detection, 238 nm.

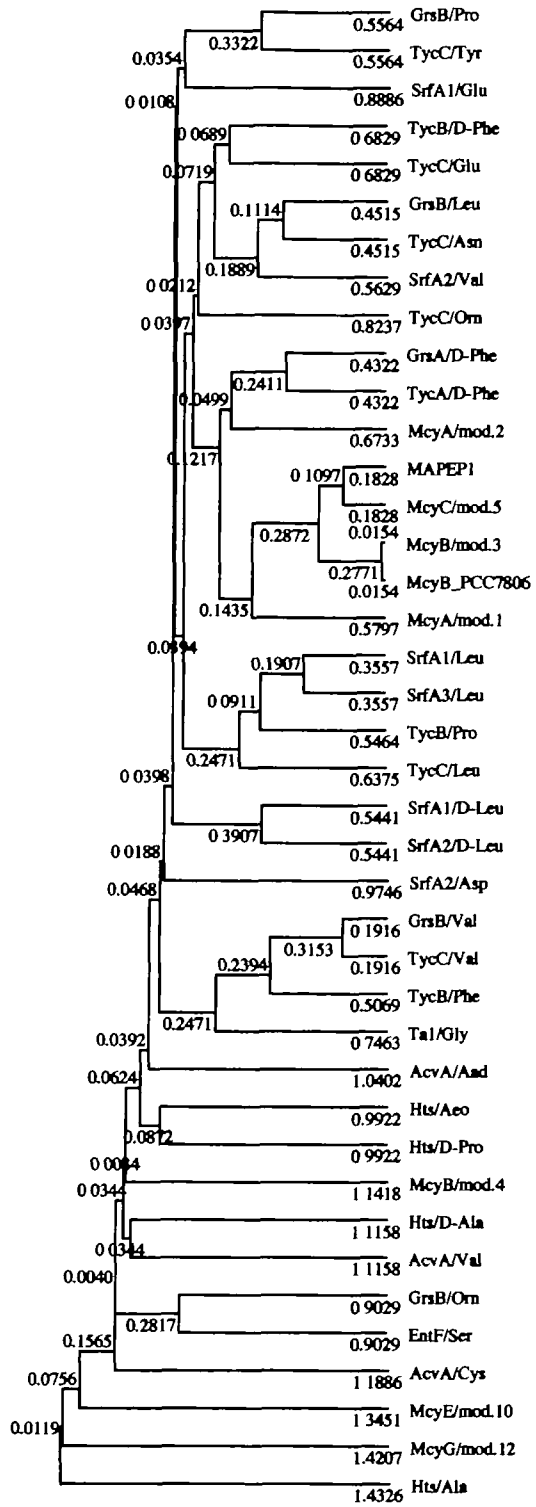

Fig. 7. **Phylogenetic tree representation of the apparent evolutionary distances of amino acid recognizing and activating domains.** GrsA and B, gramicidin S (*32*, *33*); SrfA, surfactin (*34*, *35*), Tyc, tyrocidine (*36*); Ta1, antibiotic TA (*24*); EntF, enterobactin (*39*); Mcy, microcystin (*10* and this work); McyB_PCC7806, microcystin (*9*); MAPEP1 (*48*), Acv, isopenicillin (*37*); Hts, HC-toxin (*38*).

cillin N (*37*), HC-toxin (*38*), and enterobactin (*39*) were examined by Unweighted Pair-Group Method Analysis (UPGMA) using the program contained in GENETYX-MAC (Fig. 7). The results showed that adenylation regions in the Mcy modules could be divided into two groups; one group contains modules 1, 2, 3, and 5, and other group contains modules 4, 10, and 12. Furthermore, the latter group exhibited a close relationship to those of fungi [isopenicillin N (Acv in Fig. 7) and HC-toxin (Hts)]. We also constructed

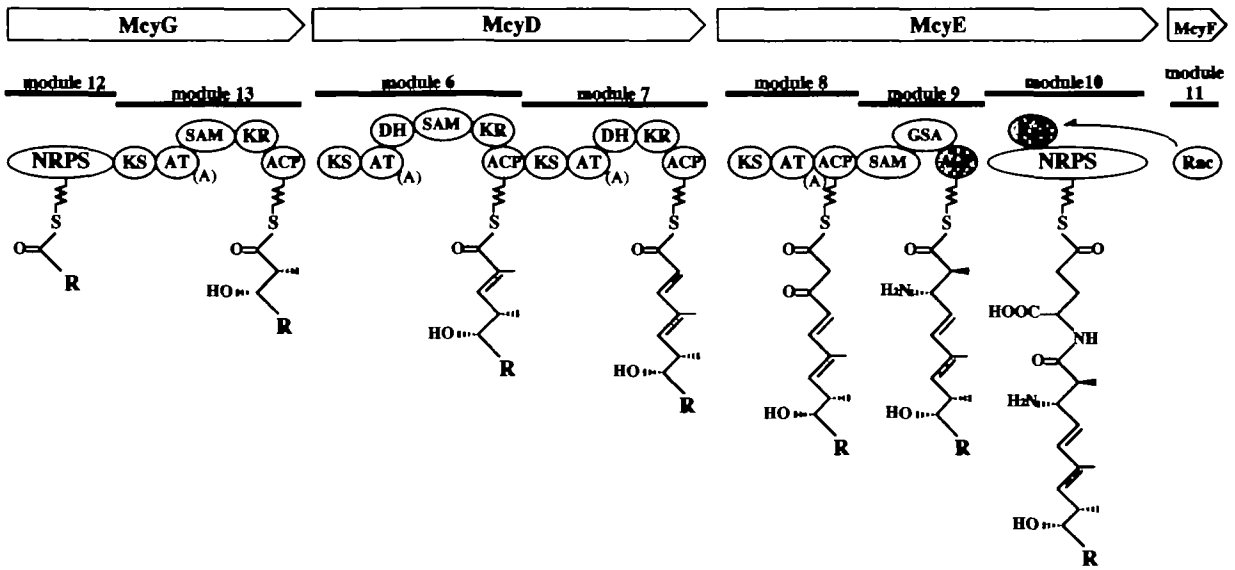

Fig. 8 **Presumed modular structure of the microcystin PKS.** (R-CO-) indicates phenylalanine or a phenylalanine-derivative as a starter unit (A) Elongation through condensation of the acetate molecule

phylogenetic trees by means of the neighbor-joining method and obtained similar results (data not shown).

DISCUSSION

In a previous study, the *orf1* gene, which showed a high degree of similarity to PKSs, was observed upstream of *mcyA* (*10*). Information about polyketide synthase genes and the genetic control of their biosynthesis in cyanobacteria is still limited. Moore *et al.* conducted [^{13}C]NMR analysis of MCYST-LR with L-amino acid, acetate and pyruvate as precursors, and suggested that the Adda molecule in microcystin is biosynthesized from the carbon skeleton of phenylalanine and four molecules of acetate (*11*).

In this study, by sequencing an additional 34 kb of DNA, we completed the characterization of the microcystin synthetase gene cluster, which is composed of seven genes including thirteen modules. Interestingly, the PKS module coupled to the NRPS module was found in McyE and McyG, and gene disruption analysis revealed that the PKS genes are involved in microcystin synthesis. Recently, Paitan *et al.* reported a PKS module coupled to a NRPS (*24*). They suggested that the polyketide antibiotic TA from *Myxococcus xanthous* is synthesized from a unique glycine-derived starter unit, which is activated by a large peptide containing a NRPS module in its N-terminal region, lying adjacent to the PKS module. In microcystin synthetase, the NRPS modules coupled to the PKS module are located in the C-terminal region of McyE and the N-terminal region of McyG, respectively (Fig. 2). According to the order of the amino acids in the microcystin molecule and the arrangement of additional domains for modification of amino acids, we presumed that McyA, B, and C are responsible for the activation of Mdha, D-Ala, L-Leu, D-MeAsp, and L-Arg (*10*). Therefore, if the *mcyD*, *E*, *F*, and *G* genes are involved in the biosynthesis of Adda, and the incorporation of Adda and Glu into the microcystin molecule, both NRPS modules are presumed to be responsible for the activation of Phe or a phenylalanine-derivative, and Glu.

In other bacterial systems such as erythromycin and rapamycin, the genes encoding the PKSs were shown to be organized in an operon (*16*). The biosynthesis of the polyketide chain is initiated by covalent loading of activated monomer units onto the first holo-ACP domain (*15*). In the case of 6-deoxyerythronolide B synthase, the starter unit was governed by the AT and ACP closest to the N-terminus of the synthase (*12*). General PKSs have an initiation domain for the starter unit with ACP or otherwise a loading domain, however, this domain is not observed in the N-terminus of McyD. The polyketide antibiotic TA is thought to be synthesized from a glycine-derived starter unit (*24*). A peptide synthetase domain which is coupled to a polyketide synthase module is likely to be involved in the processing and activation of this glycine-derived starter unit. Phylogentic analysis showed that modules 10 and 12 are unique (Fig. 7). Therefore, in the case of microcystin synthesis, a NRPS domain (module 12) at the N-terminus of McyG is likely to be involved in the processing and activation of the starter unit. In the case of the antibiotic TA synthase, an AT domain subsequent to the PKS module, which is adjacent to the NRPS module, is absent. On the other hand, PKS module 13 of microcystin synthetase contained a AT domain. To determine the substrate specificity of module 12, comparison with the amino acid activating domains of NRPSs was carried out. However, the substrate specificity of module 12 could not be determined. Polyketide metabolites are produced through the condensation of simple acid units such as propionate and acetate. In the case of *Microcystis* PKS, we could not identify the incorporated acid unit as acetate or propionate, on alignment of the sequence motif found in the AT domain (Fig. 3B). [^{13}C]NMR analysis of microcystin produced by feeding labeled substrates to microcystin-producing cells showed that the Adda molecule is biosynthesized from four molecules of acetate (*11*). Therefore, AT domains in modules 6, 7, 8, and 13 of Mcys are thought to be involved in the incorporation of acetate units.

If the *Microcystis* PKSs are responsible for synthesis of the Adda molecule, how is Adda made? The results of [^{13}C]NMR analysis suggested that Adda is biosynthesized from the carbon skeleton of phenylalanine and four molecules of acetate (*11*). A presumed schematic representation of the Adda biosynthesis is shown in Fig. 8. As the origin of the starter unit, phenylalanine or derivative of it is presumed. A phenylalanine/derivative-containing molecule, which is activated by NRPS (module 12), is used as a starter unit, and elongated through the condensation of four acetate molecules (modules 13, 6, 7, and 8). The methylation of C^2, C^6, and C^8 in the Adda molecule (Fig. 1) is catalyzed by SAM-dependent methyltransferases in modules 13, 6, and 9, and dehydration by DH in modules 6 and 7. In the case of a phenylalanine-containing molecule as a starter unit, one carbon atom should be removed from the phenylalanine molecule in the process of Adda biosynthesis. However, the module for elimination of the carbon is not located nearby. Furthermore, an amino group derived from phenylalanine should be removed or transferred. A unique module for aminotransferase is observed in module 9 of McyE, suggesting that the amino group derived from phenylalanine is transferred to C^3 in the Adda molecule. Moore *et al.* concluded that phenylacetyl-CoA is the most probable initiator for Adda biosynthesis (*11*). When a phenylacetic acid is used as the origin of the starter unit, a NRPS (module 12) activates and incorporates a phenylacetic acid, and an amino group of an unknown substrate is transferred to C^3 in the Adda molecule by an aminotransferase encoded by *mcyE*. It has been reported that phenylacetyl-CoA can be produced from phenylacetic acid in *Pseudomonas putida* (*40*, *41*). However, to the authors' knowledge, no reports have been appeared concerning the activation and incorporation of phenylacetic acid by NRPS. We can not rule out the possibility that an other phenylalanine-derivative is a starter unit. To elucidate our working hypothesis, biochemical analysis is necessary. The γ-carboxyl group of D-Glu is linked to NRPS (module 10) through a covalent interaction with a specific thiol group. Finally, the Adda molecule is transferred to the amino group of the carboxyl thioester bound D-glutamic acid. A thioesterase-like domain, which might be involved in the release and/or cyclization of the peptide chain, was observed in the C-terminal region of McyC (Fig. 2). On the other hand, this sequence motif could not be found in McyD, E, or G. Moreover, it is unclear whether the additional gene product(s), *e.g. orf3* and *orf4*, is required for microcystin synthesis.

From the data reported here as well as previously published work (*10*), it is clear that the organization of the microcystin synthetase genes is the same in *Microcystis* sp. S-70 and *M. aeruginosa* K-139, and that this gene cluster in the S-70 strain is responsible for the production of MCYST-LR, -RR, and -YR. Dittmann *et al.* reported that the disruption of *mcyB* resulted in the disappearance of plural microcystins (*9*). The following question has been raised, how does the microcystin synthetase encoded by the single gene cluster produce plural microcystins? *M. aeruginosa* K-139 produces the MCYST-LR type and *Microcystis* sp. S-70 three types of microcystin. Partial sequence data suggest that the amino acid sequences of microcystin synthetase are highly conserved (>97%) in S-70 and K-139 (*10*). However, the identity of module 3/McyB between strains S-70 and K-139 was relatively low (unpublished data). This module may be involved in the activation of various amino acids (X in the microcystin molecule shown in Fig. 1). The differences in the amino acid sequences of the modules may reflect the type of microcystin. To elucidate a working hypothesis, biochemical and genetic analyses are currently in progress.

We are grateful to Drs. K. Hori and T. Kurotsu for the information and valuable discussion.

REFERENCES

1. Carmichael, W.W. (1994) The toxins of cyanobacteria *Sci. Am.* January, 64–72
2. Watanabe, M.F., Harada, K., Carmichael, W.W., and Fujiki, H. (eds.) (1996) *Toxic Microcystis*, CRC Press, Boca Raton, FL
3. Jochimsen, E.M., Carmichael, W.W., An, J.S., Cardo, D.M., Coolson, S.T., Holmes, C.E.M., De C. Antunes, M.B., De Melo Filho, D.A., Lyra, T.M., Barret, V.S.T., Azevedo, S.M.F.O., and Jarvis, W.R. (1998) Liver failure and death after exposure to microcystins at a hemodialysis center in Brazil *N. Engl. J. Med.* **338**, 873–878
4. MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P., and Codd, G.A. (1990) Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1 and 2A from both mammals and higher plants. *FEBS Lett.* **264**, 187–192
5. Nishiwaki-Matsushima, R., Ohta, T., Nishiwaki, S., Suganuma, M., Kohyama, K., Ishikawa, T., Carmichael, W.W., and Fujiki, H. (1992) Liver tumor promotion by the cyanobacterial cyclic peptide toxin microcystin-LR *J. Cancer Res. Clin. Oncol.* **118**, 420–421
6. Konz, D., Klens, A., Schörgendorfer, K., and Marahiel, M.A. (1997) The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multimodular peptide synthetases. *Chem. Biol.* **4**, 927–937
7. Marahiel, M.A. (1997) Protein templates for the biosynthesis of peptide antibiotics. *Chem. Biol.* **4**, 561–567
8. Stachelhaus, T. and Marahiel, M.A. (1995) Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol. Lett.* **125**, 3–14
9. Dittmann, E., Neilan, B.A., Erhard, M., von Döhren, H., and Börner, T. (1997) Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC7806. *Mol. Microbiol.* **26**, 779–787
10. Nishizawa, T., Asayama, M., Fujii, K., Harada, K-I., and Shirai, M. (1999) Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp. *J. Biochem.* **126**, 520–529
11. Moore, R.E., Chen, J.L., Moore, B.S., and Patterson, G.M.L. (1991) Biosynthesis of Microcystin-LR. Origin of the carbons in the Adda and Masp units. *J. Am. Chem. Soc.* **113**, 5083–5084
12. Donadio, S. and Katz, L. (1992) Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythaea*. *Gene* **111**, 51–60
13. Malpartida, F., Hallam, S.E., Kieser, H.M., Motamedi, H., Hutchinson, C.R., Butler, M.J., Sugden, D.A., Warren, M., McKillop, C., Bailey, C.R., Humphreys, G.O., and Hopwood, D.A. (1987) Homology between Streptomyces genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes. *Nature* **325**, 818–821
14. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A., and Khosla, C. (1995) Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* **375**, 549–554
15. Cane, D.E., Walsh, C.T., and Khosla, C. (1998) Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63–68
16. Hutchinson, C.R. (1999) Microbial polyketide synthases: More and more prolific. *Proc. Natl. Acad. Sci. USA* **95**, 3336–3338

17 Shirai, M., Ohtake, A., Sano, T., Matsumoto, S., Sakamoto, T., Sato, A., Aida, T., Harada, K-I., Shimada, T., Suzuki, M., and Nakano, M. (1991) Toxicity and toxins of natural blooms and isolated of *Microcystis* spp (Cyanobacteria) and improved procedure for purification of cultures. *Appl. Environ Microbiol* **57**, 1241–1245
18 Sakamoto, T., Shirai, M., Asayama, M., Aida, T., Sato, A., Tanaka, K., Takahashi, H., and Nakano, M. (1993) Characteristics of DNA and multiple *rpoD* homologs of *Microcystis* (*Syenchocystis*) strain. *Int J Syst Bacteriol.* **43**, 844–847
19. Sambrook, J., Fritsch, E.F., and Maniatis, M. (1989) *Molecular Cloning, a Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
20. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squarres, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., and Barrell, B.G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence *Nature* **393**, 537–544
21 Schwecke, T., Aparicio, J.F., Molnár, I., König, A., Khaw, L.E., Haydock, S.F., Oliynyk, M., Caffrey, P., Cortés, J., Lester, J.B., Böhm, G.A., Staunton, J., and Leadlay, P.F. (1995) The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin *Proc. Natl Acad Sci USA* **92**, 7839–7843
22 Tang, L., Yoon, Y.J., Choi, C., and Hutchinson, C.R. (1998) Characterization of the enzymatic domains in the modular polyketide synthase involved in rifamycin B biosynthesis by *Amycolatopsis mediterranei Gene* **216**, 255–265
23. Haydock, S.F., Aparicio, J.F., Molnar, I., Schwecke, T., Ee Khaw, L., Konig, A., Marsden, A.F.A., Galloway, I.S., Staunton, J., and Leadlay, P.F. (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA.acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett* **374**, 246–248
24. Paitan, Y., Alon, G., Orr, E., Ron, E.Z., and Rosenberg, E. (1999) The first gene in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus* codes for a unique PKS module coupled to a peptide synthetase. *J Mol. Biol* **286**, 456–474
25 Haese, A., Schubert, M., Herrmann, M., and Zocher, R. (1993) Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing *N*-methyldepsipeptide formation in *Fusarium scirpi. Mol Microbiol.* **7**, 905–914
26. Grimm, B., Bull, A., and Breu, V. (1991) Structural genes of glutamate 1-semialdehyde aminotransferase for porphyrin synthesis in a cyanobacterium and *Escherichia coli Mol.Gen. Genet.* **225**, 1–10
27. de Crécy-Lagard, V., Marlière, P., and Saurin, W. (1995) Multienzymatic non ribosomal peptide biosynthesis identification of the functional domains catalysing peptide elongation and epimerisation. *CR Acad Sci* **318**, 927–936
28 Pfeifer, E., Vrancic, M.P., von Döhren, H., and Kleinkauf, H. (1995) Characterization of tyrocidine synthetase 1 (TY1) Requirement of posttranslational modification for peptide biosynthesis. *Biochemistry* **34**, 7450–7459
29. Yohda, M., Endo, I., Ohta, T., Iida, T., Maruyama, T., and Kagawa, Y. (1996) Gene for aspartate racemase from the sulfur-dependent hyperthermophilic archaeum, *Desulfurococcus* strain SY *J Biol. Chem* **271**, 22017–22021
30. Ohotake, A., Shirai, M., Aida, T., Mori, N., Harada, K-I., Matsuura, K., Suzuki, M., and Nakano, M. (1989) Toxicity of *Microcystis* species isolated from natural blooms and purification of the toxin. *Appl. Environ Microbiol* **55**, 3202–3207
31. Takahashi, I., Hayano, D., Asayama, M., Masahiro, F., Watahiki, M., and Shirai, M. (1996) Restriction barrier composed of an extracellular nuclease and restriction endonuclease in the unicellular cyanobacterium *Microcystis* sp *FEMS Microbiol. Lett* **145**, 107–111
32 Hori, K., Yamamoto, Y., Minetoki, T., Kurotsu, T., Kanda, M., Miura, S., Okamura, K., Furuyama, J., and Saito, Y. (1989) Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene. *J Biochem* **106**, 639–645
33. Turgay, K., Krause, M., and Marahiel, M.A. (1992) Four homologus domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. *Mol. Microbiol.* **6**, 529–546
34. Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. (1993) Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis. Mol Microbiol* **8**, 821–831
35 Fuma, S., Fujishima, Y., Corbell, N., D'Souza, C., Nakano, M.M., Zuber, P., and Yamane, K. (1993) Nucleotide sequence of 5′ portion of *srfA* that contains the region required for competence establishment in *Bacillus subtilis. Nucleic Acids Res.* **21**, 93–97
36 Mootz, H.D. and Marahiel, M.A. (1997) The tyrocidine biosynthesis operon of *Bacillus brevis* Complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J Bacteriol* **179**, 6843–6850
37 Smith, D.J., Earl, A.J., and Turner, G. (1990) The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillium chrysogenum* is a 421,073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. *EMBO J.* **9**, 2743–2750
38 Scott-Craig, J.S., Panaccione, D.G., Pocard, J.-A., and Walton, J.D. (1992) The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus *Cochliobolus carbonum* is encoded by a 15.7-kilobase open reading frame. *J Biol Chem* **267**, 26044–26049
39 Rusnak, F., Sakaitani, M., Drueckhammer, D., Reichert, J., and Walsh, C.T. (1991) Biosynthesis of the *Escherichia coli* siderophore enterobactin Sequence of the *entF* gene, expression and purification of EntF, and analysis of covalent phosphopantetheine *Biochemistry* **30**, 2916–2927
40 Martínez-Blanco, H., Reglero, A., Rodriguez-Aparicio, L.B., and Luengo, J.M. (1990) Purification and biochemical characterization of phenylacetyl-CoA ligase from *Pseudomonas putida. J. Biol Chem* **265**, 7084–7090
41 Minamberes, B., Maritínez-Blanco, H., Olivera, E.R., García, B., Díez, B. Barredo, J.L., Moreno, M.A., Schleissner, C., Salto, F., and Luenogo, J.M. (1996) Molecular cloning and expression in different microbes of the DNA encoding *Pseudomonas putida* U phenylacetyl-CoA ligase. *J Biol. Chem* **271**, 33531–33538
42 Harada, K., Ogawa, K., Matsuura, K., Nagai, H., Murata, H., Suzuki, M., Itezono, Y., Nakayama, N., Shirai, M., and Nakano, M. (1991) Isolation of two toxic heptapeptide microcystins from an axenic strain of *Microcystis aeruginosa* K-139 *Toxicon* **29**, 479–489
43. Simon, R. (1984) High frequency mobilization of gram-negative bacterial replicons by the *in vitro* constructed Tn5-Mob transposon *Mol Gen Genet* **196**, 413–420
44 Fernández-Moreno, M.A., Martínez, E., Boto, L., Hopwood, D.A., and Malpartida, F. (1992) Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin *J Biol Chem* **267**, 19278–19290
45. Bibb, M.J., Biro, S., Motamedi, H., Collins, J.F., and Hutchinson, C.R. (1989) Analysis of the nucleotide sequence of the *Streptomyces glaucescens tcmI* genes provides key information about the enzymology of polyketide antibiotic biosynthesis. *EMBO J.* **8**, 2727–2736
46. Weber, G., Schorgendorfer, K., Schneider-Scherzer, E., and Leitner, E. (1994) The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr Genet.* **26**, 120–125
47. Haydock, S.F., Dowson, J.A., Dhillon, N., Roberts, G.A., Cortes, J., and Leadlay, P.F. (1991) Cloning and sequence analysis of genes involved in erythromycin biosynthesis in *Saccharopolyspora erythraea*· sequence similarities between EryG and a family of *S*-adenosylmethionine-dependent methyltransferases. *Mol Gen Genet* **230**, 120–128
48 Meißner, K., Dittman, E., and Börner, T. (1996) Toxic and nontoxic strains of the cyanobacterium *Microcystis aeruginosa* contain sequences homologous to peptide synthetase genes. *FEMS Microbiol Lett* **135**, 295–303