Cyclic heptapeptide microcystin biosynthesis requires the glutamate racemase gene

Tomoyasu Nishizawa,1,3 Munehiko Asayama1 and Makoto Shirai1,2

1,2 Laboratory of Molecular Genetics, School of Agriculture, and Gene Research Center, Ibaraki University, Ami, Ibaraki 300-0393, Japan
3 The United Graduate School, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8509, Japan

Author for correspondence: Makoto Shirai. Tel: +81 298 88 8652. Fax: +81 298 88 8653. e-mail: shirai@ipc.ibaraki.ac.jp

It was demonstrated previously that the operon consisting of the non-ribosomal peptide synthetase (NRPS) gene coupled with the polyketide synthase (PKS) gene involved in cyclic heptapeptide microcystin synthesis includes two different d-amino acid synthetase genes, an epimerization domain at the 3' end of module 2, and the racemase gene mcyF. To determine the role of mcyF in microcystin synthesis, gene-disruption and complementation analyses were carried out. Insertional mutagenesis in the mcyF gene, generated by homologous recombination, abolished only microcystin synthesis, but did not influence cell growth. Furthermore, McyF supported d-Glu-independent growth of a strain of Escherichia coli defective in d-Glu synthesis. It is concluded that mcyF is the glutamic acid racemase gene involved in the synthesis of d-Glu residues in the microcystin molecule. This is the first report of the racemase in prokaryotic NRPS.

Keywords: amino acid racemase, cyanobacteria, Microcystis, peptide synthetase gene, microcystin biosynthesis

INTRODUCTION

The cyanobacteria (blue-green algae) are Gram-negative and are exceptionally large photosynthetic prokaryotes. These bacteria proliferate in eutrophic marine and freshwater habitats, resulting in the formation of water blooms. Toxic cyanobacterial water blooms are found worldwide in eutrophic lakes, ponds and dams (Carmichael, 1994; Watanabe et al., 1996). The major cause of toxicity is hepatotoxic microcystin, and the toxicity of microcystins is due to the inhibition of protein phosphatases 1 and 2A (MacKintosh et al., 1994; Watanabe et al., 1998). The major structural component of microcystin (MCYST-XZ) is cyclo(-d-Ala-X-d-MeAsp-Z-Adda-d-Glu-Mdha-), in which ‘X’ and ‘Z’ are variable D-amino acids, Mdha is N-methyldehydroalanine, d-MeAsp is d-erythro-β-methylaspartic acid, and Adda is 3-amino-9-methoxy-10-phenyl-2,6,8-trimethyl-deca-4,6-dienoic acid (Fig. 1). More than 60 structural variants of microcystins have been found (Sivonen & Jones, 1999).

Peptides including microcystins that may contain non-protein amino acids are synthesized non-ribosomally by a large multifunctional enzyme, utilizing a thio-template mechanism called non-ribosomal peptide synthetase (NRPS) (Stachelhaus & Marahiel, 1995; Marahiel, 1997). We have cloned the microcystin synthetase genes (mcyA, -B, -C, -D, -E, -F and -G) encoding seven NRPS modules and polyketide synthases (PKSs) for synthesis of the Adda molecule (Nishizawa et al., 2000). There are three NRPSs (McyA, McyB and McyC), two NRPS-PKS hybrids (McyE and McyG) and one PKS (McyD). Furthermore, in addition to the epimerase gene of module 2, a racemase gene for D-amino acid synthesis was also observed. The microcystin molecule has three D-amino acids: d-Ala, d-MeAsp and d-Glu. The same d-amino acids incorporated in non-ribosomally synthesized peptides by known prokaryotic NRPSs are synthesized by epimerization domains. Only one epimerization domain was found at the carboxy-terminal end of the thiolation domain in module 2 of McyA. Since this module is responsible for activation and incorporation of Ala, this epimerase domain must be for D-Ala synthesis.

Many D-amino acid racemases for peptidoglycan

Abbreviations: Adda, 3-amino-9-methoxy-10-phenyl-2,6,8-trimethyl-deca-4,6-dienoic acid; MCYST, microcystin; Mdha, N-methyldehydroalanine; d-MeAsp, d-erythro-β-methylaspartic acid; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; UDP-MurNAc-L-Ala-d-Glu, UDP-N-acetylmuramoyl-L-alanyl-d-glutamate.
synthesis have been reported in eubacteria (Liu et al., 1997; Malathi et al., 1999; Pucci et al., 1995). However, no racemase genes have been reported in prokaryotic NRPS. The alanine racemase genes responsible for synthesis of cyclosporin A and HC-toxin were identified in fungal NRPS (Cheng & Walton, 2000; Hoffmann et al., 1994). In this study, we showed that the mcyF gene encodes the glutamate racemase involved in microcystin synthesis.

METHODS

Bacterial strains and growth conditions. Microcystis aeruginosa K-139 (which produces 7-desmethyl-MCYST-LR and 3,7-didesmethyl-MCYST-LR) and Microcystis sp. S-70 (which produces MCYST-LR, -RR and -YR), both of which were grown under continuous illumination (2000 lx) at 30°C in CB medium (Shirai et al., 1991). Escherichia coli DH5αMCR (Cosmo Bio.) and E. coli JM109 (ToYoBo) were used as hosts for recombinant plasmids and were grown at 37°C in 2×TY broth and 2×TY agarose. E. coli strain WM335 (leu pro trp his thyA met lac gal rpsL) was used for conjugation (see Fig. 4a). Antibiotics were used for recombinants, which were grown at final concentrations of 75 µg ampicillin ml⁻¹ and 25 µg chloramphenicol ml⁻¹.

DNA manipulation, Southern blotting and sequencing. Total DNA of Microcystis cells was isolated by the method described previously (Sakamoto et al., 1993), and DNA manipulations were performed as described by Sambrook et al. (1989). Digestion of cyanobacterial DNA was separated in 0.8% agarose gels and transferred onto Hybond-NX membranes. DNA fragments (as probes) were labelled using an ECL random prime labelling kit (Amersham Pharmacia Biotech) as described previously (Nishizawa et al., 1999). Genomic Southern blotting was performed according to the manufacturer’s instructions. The nucleotide sequence was determined by dye deoxy chain termination, using an Applied Biosystems automated sequencer (model 373S) (Nishizawa et al., 1999).

Construction of a gene-disruption plasmid, and integrative conjugation of Microcystis. The plasmids for gene disruption of mcyF by homologous recombination were constructed as follows. The 2.5 kb NcoI–EcoRI fragment containing mcyF was subcloned from pMCQ3 (Nishizawa et al., 2000) by blunt-end ligation into the Smal–BamHI site of pUC119 (ToYoBo), generating pCHI621. The 1.2 kb XbaI fragment containing the CmR gene cassette from PR107XH (Nishizawa et al., 1999) and the 1.8 kb BamHI fragment containing the mob gene from pSUP5011 (Simon, 1984) were inserted by blunt-end ligation into the BamHI site (in mcyF) and the HindIII site (in the multi-cloning site) of pCHI621, respectively, generating pJXS10. The plasmids for gene disruption were introduced into Microcystis cells (by conjugation) from E. coli S17-1, and chloramphenicol-resistant (8 µg chloramphenicol ml⁻¹) conjugants were selected as described previously (Nishizawa et al., 1999).

HPLC analysis of microcystins. Microcystins were extracted from the dried cells with 5% aqueous acetic acid, purified using BondElute ODS cartridges (Varian) and analysed by HPLC as described previously (Nishizawa et al., 1999).

Bacterial complementation test. E. coli WM335, which requires d-glutamic acid for growth, was grown in Luria–Bertani (LB) medium supplemented with 50 µg d-glutamic acid ml⁻¹ and 20 µg thymine ml⁻¹. The racemase expression vector pQE-McyF was constructed as follows (see Fig. 4a). The entire mcyF gene was amplified by PCR with primers 5'-McyF/SphI (5'-CTGATGCTAGCAAAATACCGACGCTGACGGC-3') and 3'-McyF/HindIII (5'-CTCAGCTTTTTTGGGGATTACGTCGGTATACGTCGGT-3') (restriction sites are underlined and modified sequences are shown in italics). PCR was carried out on chromosomal M. aeruginosa K-139 DNA with Taq DNA polymerase (Nippon Gene). PCR products were purified, digested with SphI and HindIII, and cloned into the SphI and HindIII sites of the pQE70 vector (Qiagen). Both the 5'- and the 3'-cloning sites were confirmed by sequencing. The primers (and sequences) used were as follows: QE/F2 primer, 5'-TTGCTTTTGGAGCCCGATGATAGC-3'; QE/R1 primer, 5'-CATGAGCTTTTGGGGATTACGTCGGTATACGTCGGT-3'. HindIII digestion of pQE70 resulted in removal of the His-tag sequence and appended the amino acid sequence SLIS at the C-terminus of the recombinant protein.

Expression of McyF. E. coli WM335 cells transformed with pQE-McyF were inoculated into LB medium supplemented with ampicillin (75 µg ml⁻¹) and thymine (20 µg ml⁻¹) and were grown at 37°C with shaking for over 20 h. Expression

---

**Fig. 1.** Microcystin biosynthesis genes. The general structure of microcystin-LR (MCYST-LR) is cyclic (D-Ala-K-D-MeAsp-Z-Adda-D-Glu-Mdha-), in which X and Z are variable L-amino acids. D-MeAsp, D-erythro-β-methylaspartic acid; Adda, 3-amino-9-methoxy-10-phenyl-2,6,8-trimethyl-deca-4,6-dienoic acid; Mdha, N-methyldehydroalanine. This figure shows MCYST-LR.
was analysed by SDS-PAGE and staining with Coomassie brilliant blue according to the method of Laemmli (1970).

**Computer analysis of DNA.** The DNA sequences were assembled and analysed using GENETYX-MAC software (Software Development).

**RESULTS AND DISCUSSION**

**Characterization of the mcyF gene of M. aeruginosa**

Previously, we characterized the microcystin synthetase gene cluster and found the racemase gene, mcyF, in this cluster (Nishizawa et al., 2000). The mcyF gene is located 35 bp downstream of the PKS–NRPS hybrid gene mcyE, which contains the NRPS module at its 3′ end. The mcyF gene is 753 bp in length, encoding a 251-residue polypeptide with a predicted molecular mass of 27990 Da. The putative Shine–Dalgarno sequence (AGGAGA) was found before the putative initiation codon. The NRPS module present at the 3′ end of the mcyF gene is presumed to be involved in activation and incorporation of glutamic acid. Moore et al. (1991) demonstrated (by $^{13}$C-NMR analysis) that d-Glu in the microcystin molecule is derived from l-Glu, and that d-MeAsp does not arise from rearrangement of Glu. In the database search, McyF showed 30-4% identity with respect to the aspartate racemase of Desulfovoccus (Nishizawa et al., 2000).

d-Glutamic acid is fairly widespread in eubacteria, as d-amino acids constitute part of the fundamental tetrapeptide chain of murein in the cell wall (as common components of peptidoglycans) (Fotheringham et al., 1998; Liu et al., 1997; Malathi et al., 1999; Pucci et al., 1995). d-Glutamate in a free form is incorporated into the common peptidoglycan precursor UDP-N-acetyl-muramoyl-l-alanyl-d-glutamate (UDP-MurNac-l-Ala-d-Glu), through the action of UDP-MurNac-l-Ala-d-Glu synthetase (EC 6.3.2.9). In bacteria, enzymes are involved in the biosynthesis of d-glutamic acid through two different pathways: d-glutamate racemase and d-amino acid aminotransferase (Malathi et al., 1999). Cysteine residues at the active sites of motif 1 [[(V/I)L]x(P/A)CNTA(H/T)] and motif 2 [[(I/V)(A/L)GCTH(E/H)x(S/P)], which are thought to form the catalytic centres of these cofactor-independent racemases (Yohda et al., 1996), were also conserved in McyF (Nishizawa et al., 2000).

**Disruption of the mcyF gene**

To determine whether the cloned racemase gene, mcyF, in the mcy operon is required for microcystin biosynthesis, gene disruption of mcyF in M. aeruginosa K-139 and Microcystis sp. S-70 was carried out. However, no disruptants were isolated for strain K-139. The restriction barrier of strain K-139 may interfere with recombination (Takahashi et al., 1996). Genomic Southern hybridization demonstrated the presence of mcyF in Microcystis sp. S-70, which produces MCYST-LR, -RR and -YR (Nishizawa et al., 2000). Therefore, inserional mutagenesis of mcyF in strain S-70 was performed, and the chloramphenicol-resistant conjugant Microcystis sp. S-70/JX1 was isolated.

To demonstrate integration of the Cm$^R$ cassette into the mcyF gene on the chromosome, genomic Southern hybridization was carried out. Total DNA from mutant S-70/JX1 was digested with either HinII or HindIII and then probed with the 0·75 kb BglII fragment containing a part of the racemase gene (Fig. 2a). A 2·9 kb signal and a 5·0 kb signal, obtained with HinII and HindIII, respectively, were detected in the control (Microcystis sp. S-70). A 0·8 kb signal and a 6·2 kb signal, however, were detected in mutant JX1 (Fig. 2b). In the case of HindIII digestion, the position of this signal coincided with that of a signal obtained with the probe of the Cm$^R$ cassette fragment (data not shown). Furthermore, to confirm mcyF gene disruption and gene organization in

---

**Fig. 2.** Southern blot analysis of the mcyF-disrupted conjugant. (a) Schematic representation of insertional inactivation of the mcyF gene. The mcyF probe (black bar) is shown. Hc, HinII; H, HindIII. (b) Genomic Southern hybridization analysis. Chromosomal DNA of Microcystis sp. S-70 was digested with either HinII or HindIII. Lanes: 1, wild-type with HinII; 2, wild-type with HindIII; 3, JX1 mutant with HinII; 4, JX1 mutant with HindIII. The sizes of the signals are indicated in kb on both sides.
strain S-70JX1, the 6·2 kb HindIII fragment, including mcyE–mcyF–mcyG (Fig. 2b), was cloned from this strain. Sequence analysis showed that the gene organization in strain S-70 was the same as that in strain K-139, and nucleotide sequence identity between both strains was 95·5%.

To investigate the role of mcyF, we used HPLC to analyse microcystin production. The extracts from the control (Microcystis sp. S-70) showed three peaks for MCYST-LR, -RR and -YR (Fig. 3a). In contrast, the S-70JX1 mutant lacked all of these microcystin peaks (Fig. 3b). Moreover, the mcyF mutants showed normal growth and development on agarose plates and in liquid culture, indicating that McyF has no essential housekeeping function. These observations revealed that the mcyF gene encoding racemase is specifically involved in microcystin biosynthesis.

Genomic Southern hybridization using a mcyF probe revealed the absence of mcyF in microcystin-non-producing Microcystis strains and in Synechocystis PCC 6803 (Nishizawa et al., 2000). Two racemase genes, d-glutamate racemase and d-alanine racemase, are found in microcystin-non-producing Synechocystis PCC 6803 in the database of the Kazusa Institute, Chiba, Japan. The mcyF-disruptant exhibited significant growth similar to that of the control strain. Therefore, the mcyF gene is required only for microcystin synthesis, but not for fundamental growth.

**Fig. 3.** HPLC of methanol extracts of Microcystis sp. S-70 cells. Microcystins were extracted from dried cells of the wild-type (a) and the mutant JX1 (b). Peaks: RR, MCYST-RR; YR, MCYST-YR; LR, MCYST-LR. Column, Cosmosil 5C18ARI (150×4·6 mm i.d.); mobile phase, CH3CN/0·01 M trifluoroacetic acid (30:70, v/v); flow rate, 1·0 ml min−1; detection, 238 nm.

Complementation of a bacterial mutant deficient in D-Glu synthesis

To confirm that the mcyF gene encodes the glutamate racemase, the McyF racemase was expressed in d-glutamic-acid-requiring E. coli WM335, using the T5 promoter transcription–translation system of pQE vectors (Fig. 4a). E. coli WM335 transformed with pQE-McyF expressed a protein of about 28 kDa (Fig. 4b), the molecular mass of which coincided well with the value predicted for McyF on the basis of its nucleotide sequence. The pQE-McyF-transformed E. coli grew exponentially and survived on agarose plates and in LB liquid medium (data not shown) without d-glutamic acid as well as in those supplemented with 50 µg d-glutamic acid ml−1. However, E. coli WM335 carrying the expression vector pQE70 or pCHI621 showed no
growth on agarose plates without \(d\)-glutamic acid, and a rapid decrease in \(OD_{600}\) was observed shortly after inoculation into liquid medium without \(d\)-glutamic acid (data not shown). These observations showed that the mcyF gene was capable of complementation of the \(d\)-glutamic-acid-requiring phenotype in \(E.\ coli\). Gene disruption and complementation analysis demonstrated that \(d\)-Glu in the cyclic heptapeptide microcystin is synthesized from the primary metabolite \(L\)-Glu by a glutamate racemase (McyF). These results indicated that McyF is involved in \(d\)-glutamic acid synthesis. Moreover, this is the first report showing the occurrence of a glutamate racemase gene in prokaryotic non-ribosomal peptide synthesis.

**Phylogenetic analysis of the \(d\)-amino acid racemase**

The microcystin molecule is composed of the following three \(d\)-amino acid residues: \(d\)-Ala, \(d\)-MeAsp and \(d\)-Glu (Fig. 1). The microcystin synthetase was shown previously to contain one epimerization domain in McyA and one racemase, McyF (Nishizawa et al., 2000). In NRPS, \(d\)-amino acid synthesis is generally catalysed by an epimerase, which is encoded by an epimerization
domain at the carboxy-terminal end of the thiolation domain (Stein et al., 1995). Our results indicated that a glutamic acid racemase, McyF, is responsible for the production of d-Glu.

On the other hand, a sequence similarity search of d-amino acid racemases from bacteria and fungi, which are available in the GenBank, EMBL and DDBJ databases, was carried out using the program in GENETYX-MAC (see Methods). Our results revealed that McyF has 22.4–30.1% sequence similarity to aspartate racemase, 16.9–23.7% similarity to glutamate racemase, and 15.1–21.2% similarity to alanine racemase. The phylogenetic relationships between these racemase genes were analyzed by the neighbour-joining method, using GENETYX-MAC, and the resulting dendrogram is shown in Fig. 5. These racemase genes were tightly clustered, except for the glutamate racemase of Haemophilus influenzae. d-Amino acid racemases can be roughly divided into two groups, namely the Ala type and the Glu type. Aspartate racemases belong to the glutamate racemase cluster. Two cysteine residues, which are thought to form the catalytic centre, were highly conserved among the aspartate and glutamate racemases, as well as among their surrounding amino acid sequences (Nishizawa et al., 2000).

The only racemase previously reported to be involved in non-ribosomal peptide synthesis was the alanine racemase in fungi (Cheng & Walton, 2000; Hoffmann et al., 1994). The involvement of specific racemases in prokaryotic non-ribosomal peptide synthesis has not been reported. Their racemase genes are unrelated, at the primary amino acid level, to any known bacterial alanine racemases (Cheng & Walton, 2000; Hoffmann et al., 1994). This analysis indicated that McyF belongs to the glutamic acid racemase group. Microcystin synthetase, including both types of d-amino acid synthesis enzymes, epimerase and racemase, is a unique NRPS. Interestingly, McyF is grouped with aspartate racemases in the phylogenetic tree (Fig. 5). Since the microcystin molecule includes d-MeAsp, it is possible that McyF is responsible for d-MeAsp synthesis together with d-Glu synthesis. However, this possibility has been excluded by the work of Moore et al. (1991), who concluded from the [1,2-13C]acetate-labelling pattern that MeAsp is synthesized from acetyl-CoA and pyruvic acid.

ACKNOWLEDGEMENTS

We thank Drs K.-I. Harada and K. Fujii for their help with HPLC analysis. We are grateful to Drs W. Messer and M. Wachi for gifts of the strains used in this study.

REFERENCES


Received 3 October 2000; revised 6 December 2000; accepted 29 January 2001.