The *mcyF* gene of the microcystin biosynthetic gene cluster from *Microcystis aeruginosa* encodes an aspartate racemase

Heike SIELAFF*, Elke DITTMANN*, Nicole TANDEAU DE MARSAC†, Christiane BOUCHIER‡, Hans VON DÖHREN§, Thomas BÖRNER*[1] and Torsten SCHWECKE§

*Institut für Biologie, Abteilung Genetik, Humboldt Universität zu Berlin, Chausseestrasse 117, D-10115 Berlin, Germany, †Unité des Cyanobactéries, Département de Microbiologie Fondamentale et Médicale, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France, ‡Génopole Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France, and §Institut für Chemie, Abteilung Biochemie und Molekulare Biologie, Technische Universität Berlin, Franklinstrasse 29, D-10587 Berlin, Germany

Microcystins are hepatotoxic, non-ribosomal peptides produced by several genera of freshwater cyanobacteria. Among other enzymic activities, in particular those of peptide synthetases and polyketide synthases, microcystin biosynthesis requires racemases that provide D-aspartate and D-glutamate. Here, we report on the cloning, expression and characterization of an open reading frame, *mcyF*, that is part of the *mcy* gene cluster involved in microcystin biosynthesis in the *Microcystis aeruginosa* strain PCC 7806. Conserved amino acid sequence motifs suggest a function of the McyF protein as an aspartate racemase. Heterologous expression of *mcyF* in the unicellular cyanobacterium *Synechocystis* PCC 6803 yielded an active His_6-tagged protein that was purified to homogeneity by Ni^{2+}-nitriloacetate affinity chromatography. The purified recombinant protein racemized in a pyridoxal-5′-phosphate-independent manner L-aspartate, but not L-glutamate. Furthermore, we have identified a putative glutamate racemase gene that is located outside the *mcy* gene cluster in the *M. aeruginosa* PCC 7806 genome. Whereas homologues of this glutamate racemase gene are present in all the *Microcystis* strains examined, *mcyF* could only be detected in microcystin-producing strains.

Key words: aspartate racemase, cyanobacteria, glutamate racemase, microcystin, *Microcystis aeruginosa*, non-ribosomal peptide.

INTRODUCTION

Microcystins are a family of approximately 60 monocyclic heptapeptides produced by many, but not all, strains of different cyanobacterial genera; in particular, *Microcystis*, *Nostoc*, *Oscillatoria* and *Anabaena* [1]. They exhibit strong hepatotoxicity, at least in part due to inhibition of eukaryotic serine/threonine protein phosphatases [2]. They have been suggested further to promote primary liver tumours [3] and colorectal cancer [4]. Thus the frequently observed massive development of microcystin-producing strains in water bodies represents a potential health hazard [1].

All microcystins contain several unusual amino acid residues and differ structurally, primarily in two positions of their backbone. The common structure is cyclo(Adda-D-Glu-Mdha-D-Ala-L-X-D-Asp/D-MeAsp-L-Z), where 'X' and 'Z' are variable L-amino acids, 'Adda' is 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoic acid, 'D-MeAsp' is 3-methylaspartic acid and 'Mdha' is *N*-methyldehydroalanine (Figure 1A). The most common derivative is microcystin-LR, in which amino acids at positions X and Z are leucine (L) and arginine (R) respectively. The D-aspartate residue can be methylated at the β-carbon atom (giving rise to MeAsp). The biosynthetic gene clusters governing microcystin biosynthesis in two strains of *Microcystis aeruginosa*, PCC 7806 [5] and K-139 [6,7], have been cloned and sequenced. Both clusters are virtually identical and consist of two divergently transcribed operons encoding a hybrid type I polyketide synthase (PKS)/non-ribosomal peptide synthetase (NRPS) system and modifying enzymes (Figure 1B). Most of the activities necessary for the biosynthesis of the heptapeptide backbone could be readily inferred from the DNA sequence, and yet, although microcystins contain three D-amino acids (D-glutamate, D-alanine and D-MeAsp/D-aspartate), there are only two obvious racemase activities present in the gene cluster. One is an integral epimerase domain of an NRPS module responsible for incorporating D-alanine into the peptide backbone. Since the modules responsible for the incorporation of D-glutamate and D-MeAsp/D-aspartate into the peptide backbone do not contain an integrated epimerase domain, the two respective amino acids must be converted by stand-alone racemases either before or after incorporation into the peptide backbone. A putative racemase gene, *mcyF*, is present in the cluster and is located between two hybrid *PKS/NRPS* genes. Disruption of this gene abolished microcystin production in *M. aeruginosa* K-139 and has been reported to be responsible for the formation of D-glutamate [8]. We have undertaken a detailed *in silico* analysis of the functional organization of the deduced amino acid sequence of *mcyF*, and report here that the gene shows a higher similarity to authentic aspartate racemases than to glutamate racemases. We purified a recombinant version of this protein and characterized its substrate specificity. We provide evidence for L-aspartate being the substrate for the racemase encoded by *mcyF*. Moreover, we detected a gene for a putative L-glutamate-specific racemase located outside the *mcy* gene cluster that may be responsible for the formation of the D-glutamate incorporated at position 6 of microcystin in *M. aeruginosa*.

EXPERIMENTAL

Materials

L-[U-^{14}C]aspartate (7.4 TBq/mol) and L-[U-^{14}C]glutamate (9.25 TBq/mol) were purchased from NEN Life Science Products (Boston, MA, U.S.A.). All other materials were reagent grade.

Abbreviations used: D-MeAsp, 3-methylaspartic acid; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; PLP, pyridoxal 5′-phosphate.

[1] To whom correspondence should be addressed (e-mail thomas.boerner@rz.hu-berlin.de).

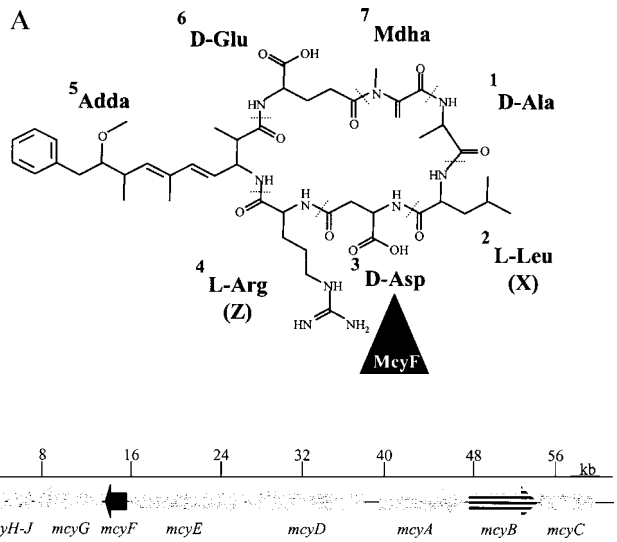

Figure 1 Structure of microcystin-LR (A) and the biosynthetic gene cluster for microcystin from *Microcystis aeruginosa* PCC 7806 (B)

(**A**) The structure of microcystin-LR is shown. The black triangle representing McyF points towards its product D-aspartate incorporated at position 3 of microcystin. For further explanation of the nomenclature describing the other positions, see the Introduction. (**B**) The biosynthetic gene cluster for microcystin from *Microcystis aeruginosa* PCC 7806 is shown. Arrows represent the genes. For a description of the complete functionalities of all open reading frames, refer to [5]. *McyF* is shown in black. D-Aspartate/D-MeAsp are thought to be activated as adenylates, and to be incorporated into the final product microcystin by the non-ribosomal peptide synthetase McyB, which is shown in black-and-white stripes.

Bacterial strains, plasmids and growth conditions

Synechocystis PCC 6803 wild-type and transconjugants, as well as *M. aeruginosa* strain PCC 7806, were grown as batch cultures in liquid BG-11 medium [9]. For growth on plates, 1 % (w/v) agar was added. Cultures were grown under continuous aeration and constant illumination at 29 °C. Light conditions were as follows: low-intensity light (8 $\mu E \cdot m^{-2} \cdot s^{-1}$), medium-intensity light (50 $\mu E \cdot m^{-2} \cdot s^{-1}$) and high-intensity light (165 $\mu E \cdot m^{-2} \cdot s^{-1}$). Light intensities were measured using a model LI-250 light-meter (LI-COR, Frankfurt, Germany). Transconjugants were grown in the presence of 2 μg/ml gentamicin. *Escherichia coli* strain XL1 Blue (Stratagene, La Jolla, CA, U.S.A.), used as the host for plasmid construction, was grown in Luria–Bertani medium at 37 °C [10]. Recombinant *E. coli* strains were grown in the presence of 75 μg/ml ampicillin or 5 μg/ml gentamicin.

Bioinformatics

Sequence analysis was performed using the MULTALIN (http://prodes.toulouse.inra.fr/multalin/) and GENEBEE (http://www.genebee.msu.ru) software packages, and the BLAST programs at The National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) and The European Bioinformatics Institute (http://www.ebi.ac.uk). A partial *M. aeruginosa* genome database was screened using the TBLASTN facility provided (Institut Pasteur, Paris, France).

Construction of plasmids

Genomic DNA of *Microcystis* sp. was isolated essentially as described by Franche and Damerval [11]. The *mcyF* gene was amplified by PCR using the following primers: 5′-***CATATG***CATCATCATCATCATAAGACAAAACTACCGATTCTTGG-3′ (mcy-FR4; forward) and 5′-GTACCTGATA***CATATG***CGATGGTC-3′ (mcy-FR3; reverse). (Bold letters show the restriction sites; modified sequences are in italics. Bases encoding the N-terminal His_6-tag are shown underlined.)

Glutamate racemase genes were amplified by PCR using the following primers: 5′-TAATACCAGTTCCGCCTTGG-3′ (gluF) and 5′-GGTAGGGATGTTTCGGGATT-3′ (gluR). PCR was performed in a final volume of 50 μl using *M. aeruginosa* genomic DNA as template and Advantage 2 polymerase mix (BD Biosciences Clontech, Palo Alto, CA, U.S.A.). Conditions were as follows: denaturation at 95 °C for 2 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 57 °C for 30 s, extension at 72 °C for 1 min, followed by a final extension step at 72 °C for 10 min. The *mcyF* PCR product was gel-purified and cloned into the pGEM-T-vector (Promega, Madison, WI, U.S.A.). The sequence was verified on both strands on an ABI 373 automatic DNA sequencer (Applied Biosystems Inc., Foster City, CA, U.S.A.). For expression in *E. coli*, the 780 bp *Nde*I fragment encoding the modified version of *mcyF* was subcloned into pet15b vector (Novagen, Madison, WI, U.S.A.); for expression in *Synechocystis* PCC 6803, it was subcloned in the vector pPbs [12], a pGEM-T-vector derivative carrying the light-inducible *psbA*2 promoter from *Synechocystis* PCC 6803. A 1240 bp *Xho*I/*Mlu*I fragment of the resulting plasmid was inserted into expression vector pVZ322 [13], which was linearized with *Xho*I and *Mlu*I. This construct was transferred to *Synechocystis* PCC 6803 cells by conjugation, as described by Elhai and Wolk [14].

Purification of recombinant protein from *Synechocystis* PCC 6803[pVZ322*mcyF*]

BG-11 medium (1 litre) supplemented with 2 μg/ml gentamicin was inoculated with a seed culture to give an attenuance (D_{750}) of 0.2. Cells were harvested at $D_{750} = 0.8$–1.0 by centrifugation for 10 min at 5000 ***g***. A typical culture yielded between 200 and 300 mg of cells (dry mass). Cell pellets were resuspended in lysis buffer [50 mM NaH_2PO_4 (pH 8.0)/300 mM NaCl/10 mM imidazole/0.5 mM dithiothreitol/100 μM PMSF/100 μM benzamidine] and broken with glass beads (0.17–0.18 mm and 0.11 mm; Braun Biotech International, Melsungen, Germany) in a Mixer Mill MM2 (Retsch GmbH, Haan, Germany) for 40 min at 4 °C. Beads and cell debris were removed by centrifugation at 18 000 ***g*** for 40 min at 4 °C. The remaining pellet was washed twice with lysis buffer, and supernatants were combined and centrifuged at 18 000 ***g*** for 40 min at 4 °C. The resulting supernatant was applied to 1 ml of a Ni^{2+}-nitrilotriacetic acid matrix (Qiagen, Hilden, Germany). The column was washed with 3×4 ml of lysis buffer containing 20 mM imidazole, and bound protein was eluted from the column with a gradient of 50–250 mM imidazole in lysis buffer. McyF was eluted at an imidazole concentration of approx. 150–250 mM. To reduce the imidazole concentration, McyF containing fractions were desalted on a PD 10 column (Pharmacia GmbH, Erlangen, Germany) equilibrated with lysis buffer. Protein concentrations were determined by the dye-binding method of Bradford [15] with BSA as a standard. Proteins were separated on SDS/15 % polyacrylamide gels and blotted on to Hybond C Extra membrane (Amersham Biosciences Europe GmbH, Freiburg, Germany). Protein blotting was performed as described by Lamparter et al. [16]. Protein gels were stained with Coomassie Brilliant Blue R-250. Polyclonal antibodies against McyF were kindly provided by Dr Michael Hisbergues (Department of Microbiology of Ecosystems, Institute for Biology, Lille, France).

Antigen–antibody complexes were detected using anti-rabbit IgG conjugated to horseradish peroxidase, and then visualized using SuperSignal West Pico chemiluminescent substrate (Pierce Biotechnology, Rockford, IL, U.S.A.).

Racemase activity assays and separation of amino acid enantiomers on chiral TLC

Activities were determined as follows: all incubations contained 250 μl of enzyme solution, 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 10 mM imidazole, 20 mM dithiothreitol, 100 μM PMSF, 100 μM benzamidine and 20 mM of the specific L-amino acid in a final volume of 300 μl. The reaction mixtures were incubated at 23 °C for various periods of time. When appropriate, incubation assays were supplemented with U-^{14}C-radiolabelled L-amino acids (0.5 μCi). The reaction was stopped by addition of 2.5 vol. of acetone. Following centrifugation, the protein-free supernatant was dried under a vacuum. The remaining pellet was dissolved in 50% (v/v) methanol, and the enantiomers were separated by TLC on chiral TLC plates (Macherey-Nagel GmbH, Düren, Germany) using an acetonitrile/methanol/water (4:1:1, by vol.) mixture as the mobile phase. The TLC plates were heated for 15 min at 100 °C immediately before use. Unlabelled amino acids were detected by ninhydrin [0.3% (v/v) in acetone]. Radiolabelled amino acids were detected by exposing a PhosphoImager screen to the dried silica plates. Data were collected employing a PhosphoImager System (Bio-Rad Laboratories GmbH, München, Germany), and analysed using the Personal FX software provided.

RESULTS AND DISCUSSION

McyF sequence analysis

McyF has recently been described as a glutamate racemase [8]. A comparison of the amino acid sequence deduced from the *mcyF* gene with those of authentic racemases reveals the presence of two highly conserved cysteine residues characteristic for the pyridoxal-5′-phosphate (PLP)-independent glutamate/aspartate racemase family. Both cysteine residues have a role in catalysis by acting as bases for the de- and re-protonation of the substrate amino acid [17,18]. However, despite a generally high level of overall sequence similarity between the PLP-independent aspartate and glutamate racemases, there are considerable differences at the primary structure level that allow a clear assignment. In addition to the two active cysteine residues, glutamate racemases possess more strictly conserved residues, four of which (Asp-10, Asp-36, Glu-152 and His-186) have been shown to be crucial for activity in the *Lactobacillus fermenti* MurI glutamate racemase [19]. These amino acids are not conserved in McyF. Instead, the enzyme contains several residues, including Thr-19, Thr-89, His-91, Leu-126, Thr-129, Ile-192, Glu-197 and Asp-217, which are highly conserved in aspartate racemases. Two additional residues, Phe-137 and Ile-69, are conservative changes corresponding to tyrosine and leucine at the respective positions (Figure 2).

Hence the sequence comparison clearly suggests that McyF is involved in the racemization of aspartate, rather than glutamate. Interestingly, on the basis of the primary structural alignment, two subgroups of aspartate racemases can be defined (Figure 3). It is, as yet, unclear whether the differences between these two subfamilies at the primary sequence level are reflected by their biochemical function, because most enzymes have not been characterized enzymically. A remarkable difference between McyF and authentic aspartate racemases is the presence of Cys-88 instead of the strictly conserved Asn-88 next to the active site Cys-87. Whereas the enzyme is evidently grouped with one subgroup of the aspartate racemase family, it is less closely related than the other members of this group. This may be due to a broader substrate specificity, because the enzyme may be responsible for the racemization of L-aspartate as well as that of L-MeAsp.

Heterologous expression of *mcyF*

In order to prove experimentally our hypothesis on the basis of sequence comparisons that *mcyF* encodes an aspartate racemase, we have constructed a plasmid, pet15b-*mcyF*, for the expression of a His_6-tagged version of McyF in *E. coli*. Overexpression of *mcyF* in *E. coli* led to the formation of inclusion bodies, and did not yield enough soluble protein for biochemical analyses (results not shown). Therefore we decided to use an alternative host system. The unicellular cyanobacterium *Synechocystis* PCC 6803 has recently been used for heterologous expression [20]. We have constructed another plasmid in which a modified version of *mcyF* encoding an N-terminally His_6-tagged protein was under the control of the light-inducible *Synechocystis psbA2* promoter [12] (Figure 4). Heterologous expression of the gene in *Synechocystis* PCC 6803 gave rise to soluble McyF protein. Even though, under low-intensity light conditions ($8\ \mu E \cdot m^{-2} \cdot s^{-1}$), protein production was approx. 15 times higher than the amount produced by the wild-type *M. aeruginosa* strain, high-intensity light conditions ($165\ \mu E \cdot m^{-2} \cdot s^{-1}$) increased the yield further by approx. 2-fold, as judged from immunoblotting experiments (Figure 5A).

Purification and analysis of substrate specificity of recombinant McyF

Cells were harvested by centrifugation and broken in a cell disrupter. After centrifugation, the clear supernatant (crude extract) was applied to a Ni^{2+}-nitriloacetate column, and the target protein was purified. Figure 5(B) shows the SDS/PAGE analysis of both the crude extract and the purified McyF. The apparent M_r of the recombinant protein on SDS/PAGE is in agreement with the expected value of 28 000.

Crude protein extracts, as well as purified McyF preparations, were assayed for their ability to racemize various L-amino acids. Incubations were performed using radiolabelled and unlabelled L-amino acids as substrates. Analysis of the supernatant by TLC clearly showed that an overnight incubation at 23 °C of both a crude extract of the transconjugant strain *Synechocystis* PCC 6803[pVZ322*mcyF*] and the purified protein McyF resulted in a racemization of the substrate amino acid L-aspartate. As expected, the enzyme did not require the addition of PLP as a cofactor. When L-glutamate was used as a substrate, no formation of D-glutamate was observed under otherwise identical conditions (Figure 6).

Our experimental results thus revealed aspartate, but not glutamate, racemase activity of McyF. Nishizawa et al. [8] have shown that *mcyF*, when expressed in an *E. coli* strain defective in D-glutamate synthesis, supported D-glutamate-independent growth, suggesting that the encoded enzyme is in fact capable of racemizing L-glutamate. This observation seems to contradict our observation of aspartate racemase activity. McyF might have a dual function as both an aspartate and a glutamate racemase. However, this is unlikely, since we did not detect any glutamate racemase activity of McyF in our assays. Furthermore, according to sequence alignments, McyF clearly belongs to the aspartate branch of the phylogenetic tree of aspartate/glutamate racemases. In addition, McyF has none of the four conserved amino acid residues characteristic for glutamate racemases.

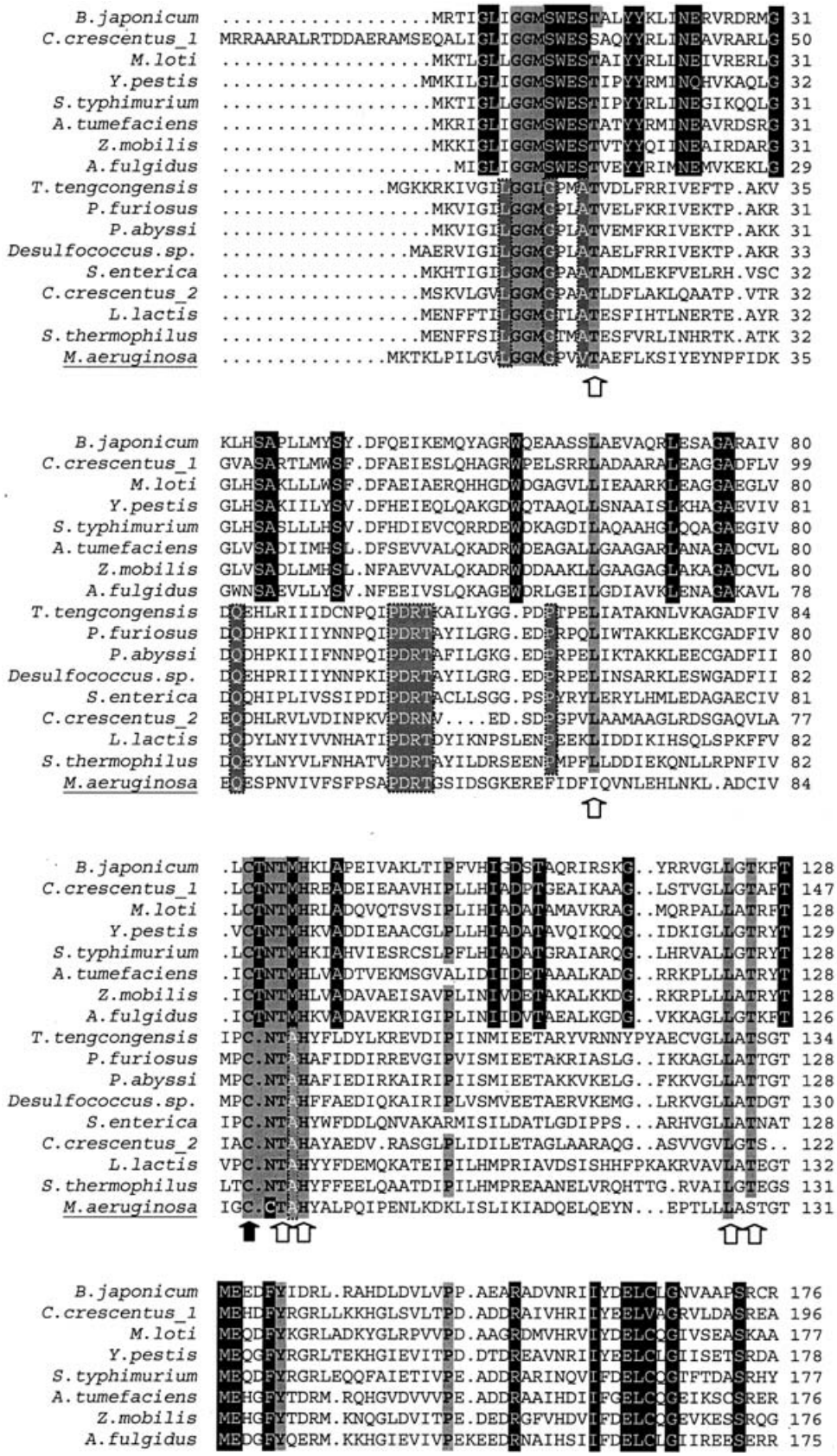

Figure 2 **For legend see facing page**

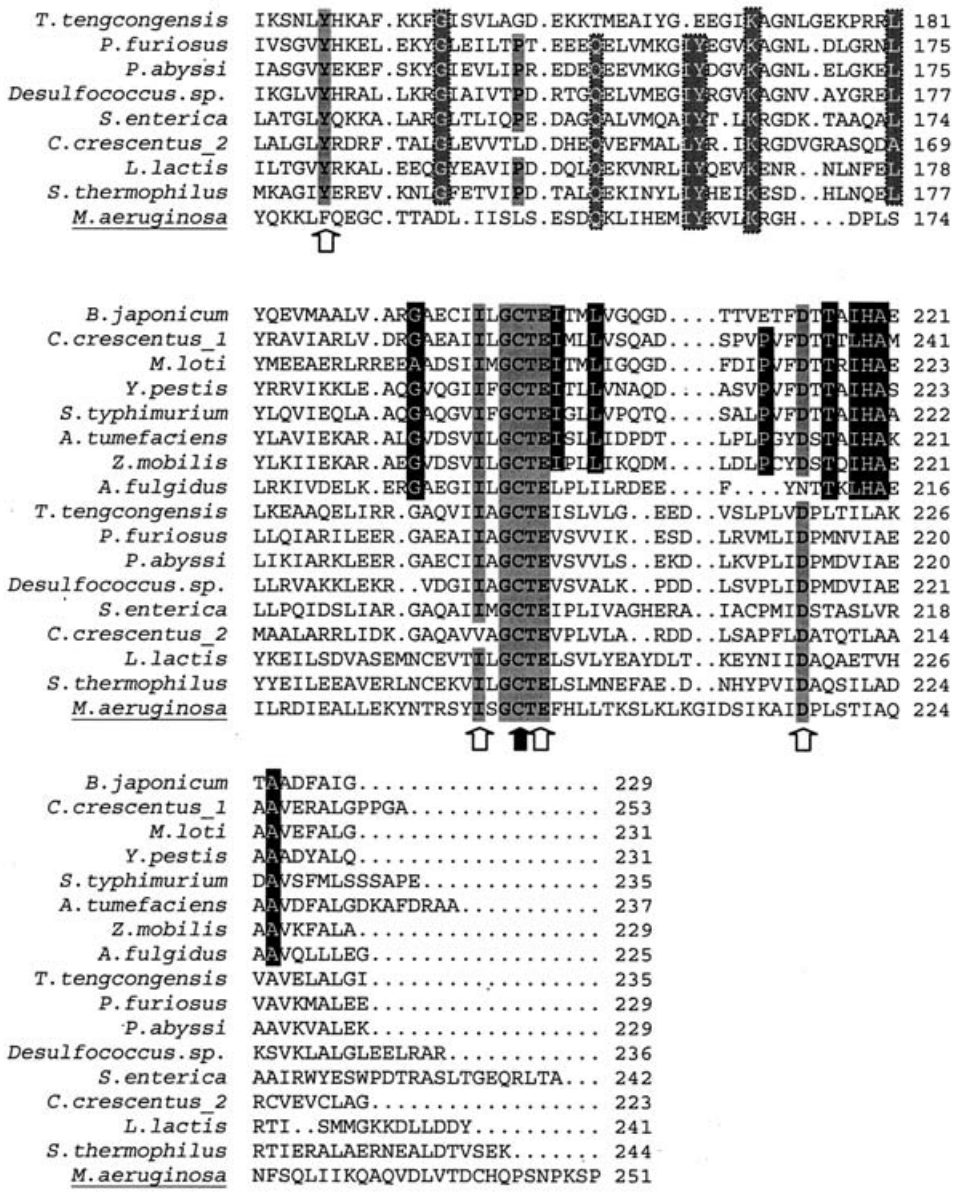

Figure 2 Sequence alignment of authentic aspartate racemases from several micro-organisms and McyF from *M. aeruginosa*

The aspartate racemases shown here are grouped into two subfamilies. Residues conserved in all aspartate racemases are shown in bold letters on grey background. Residues conserved in only one subfamily are indicated as white letters on grey and black backgrounds respectively. White arrows indicate residues that are conserved only in aspartate, and not glutamate, racemases. Black arrows indicate the two active-site cysteine residues. For an explanation of full names and accession numbers, refer to the legend to Figure 3.

Moreover, it has been shown for an aspartate racemase from *Streptococcus thermophilus* that the enzyme has a narrow substrate specificity, and does not utilize glutamate as a substrate [21]. Likewise, glutamate racemase from *Bacillus subtilis* does not accept aspartate as a substrate or as an inhibitor [22]. An explanation for the discrepancy between our results and those of Nishizawa et al. [8] might be provided by the fact that the UDP–*N*-acetylmuramoylalanyl-D-glutamate ligase ('MurD'), an enzyme catalysing the addition of D-glutamate to UDP–*N*-acetylmuramoyl-L-alanine in *E. coli* peptidoglycan biosynthesis, is inhibited by D-aspartate [23]. It is therefore tempting to speculate that this enzyme could actually use D-aspartate as a substrate. The replacement of D-glutamate with D-aspartate in peptidoglycan synthesis could explain why McyF supports growth in an *E. coli* strain defective in D-glutamate synthesis.

Distribution of genes encoding aspartate and glutamate racemases in hepatotoxic and non-toxic *Microcystis* strains

If McyF provides the D-aspartate for microcystin biosynthesis, another enzyme is needed to supply the D-glutamate that is incorporated at amino acid position 6 of microcystin. It is conceivable that *M. aeruginosa*, like other bacteria, possesses a glutamate racemase for the provision of D-glutamate for use in peptidoglycan biosynthesis. As there is no obvious candidate gene in the *mcy* gene cluster, a TBLASTN search of the *M. aeruginosa* strain PCC 7806 genome database at the Institut Pasteur (N. Tandeau de Marsac, C. Bouchier, E. Dittmann and T. Börner, unpublished work) was performed using the glutamate racemase of the cyanobacterium *Synechocystis* PCC 6803 as a query sequence. The gene thus identified is located immediately

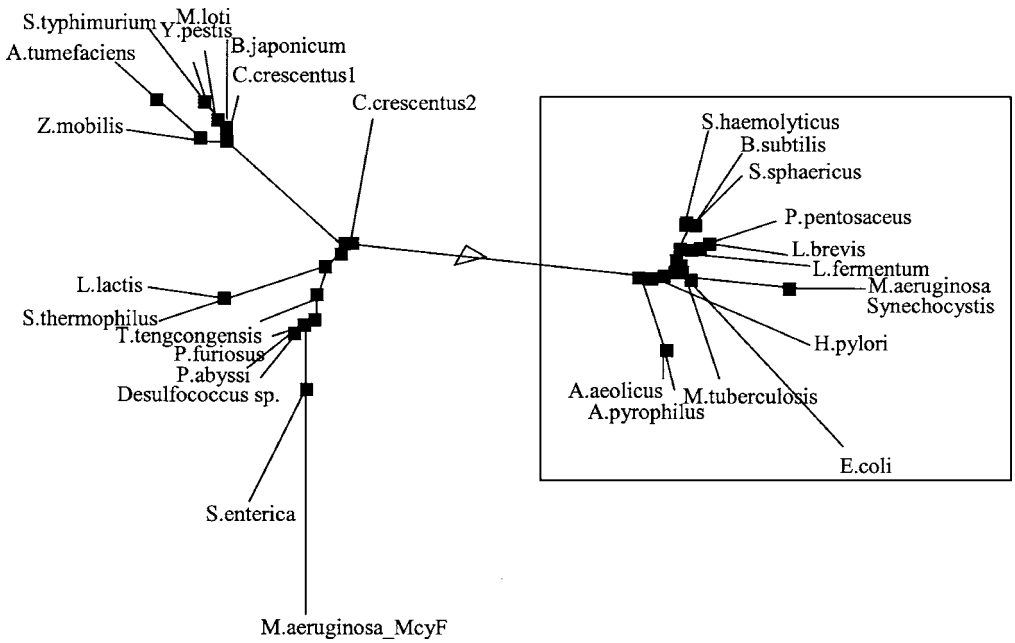

Figure 3 Radial tree of glutamate and aspartate racemases

A phylogenetic tree was constructed as described in the Experimental section from the deduced amino acid sequences of aspartate and glutamate racemases of *M. aeruginosa* and other organisms (accession numbers are shown in parentheses). The glutamate branch is boxed. The branch length is proportional to the phylogenetic distances (scale bar not shown). Aspartate racemases: *Bradyrhizobium japonicum* (BAC53419), *Caulobacter crescentus*_1 (D87655), *Mesorhizobium loti* (BAB50815), *Yersinia pestis* (AF0336), *Salmonella typhimurium* (AAL21891), *Agrobacterium tumefaciens* (AAL44882), *Zymomonas mobilis* (AAD54635), *Archaeoglobus fulgidus* (AAB89824), *Thermoanaerobacter tengcongensis* (AAM24924), *Pyrococcus furiosus* (AAL81104), *Microcystis aeruginosa* McyF (Q9RNB4), *Pyrococcus abyssi* (D75048), *Desulfurococcus* sp. (T08467), *Salmonella enterica* (AD1067), *Caulobacter crescentus*_2 (D87501), *Lactococcus lactis* (B86903), *Streptococcus thermophilus* (P29079); glutamate racemases: *Lactobacillus fermentum* (A49473), *Aquifex aeolicus* (B70329), *Aquifex pyrophilus* (P56868), *Helicobacter pylori* (NP_207344), *Lactobacillus brevis* (P48797), *Pediococcus pentosaceus* (Q08783), *Bacillus sphaericus* (P52972), *Staphylococcus haemolyticus* (P52974), *Bacillus subtilis* (CAA99552), *Mycobacterium tuberculosis* (Q10626), *Microcystis aeruginosa* PCC 7806 (the present study; AJ549183), *Synechocystis* PCC 6803 (S74824).

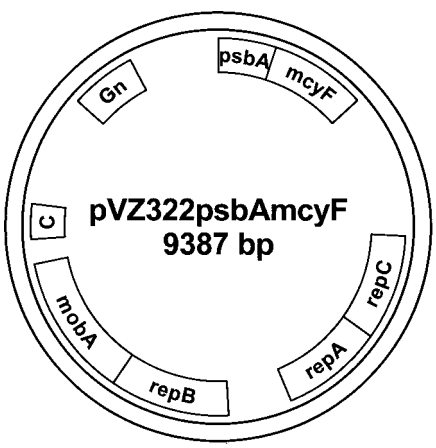

Figure 4 Diagrammatic representation of the expression vector pVZ322*mcyF*

The vector was constructed as described in the Experimental section. The *mcyF* gene is under control of the light-inducible *psbA2* promoter. Abbreviations (reading from 12 o'clock anticlockwise) are as follows: psbA, *psbA*2 promoter, Gn, gentamycin-3′-acetyltransferase, C, mobilization protein C; mobA, mobilization protein A; repB, replication protein B, repA, replication protein A; repC, replication protein C.

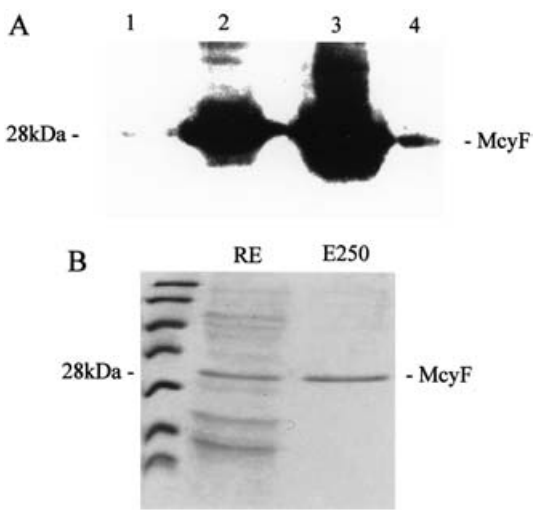

Figure 5 Immunoblot showing the relative expression levels of *mcyF* (A) and SDS/PAGE of a single-step purification of the His_6-tagged McyF protein from crude extract of *Synechocystis* PCC 6803 transformed with the expression vector pVZ322*mcyF* (B)

(**A**) An immunoblot showing the relative expression levels of *mcyF* is shown. Lane 1, *Synechocystis* PCC 6803 wild-type; lane 2, transconjugant grown under low-intensity-light conditions; lane 3, transconjugant grown under high-intensity-light conditions; lane 4 contains *M. aeruginosa* PCC 7806 wild-type. (**B**) Shown is SDS/PAGE of a single-step purification of the His_6-tagged McyF protein from crude extract of *Synechocystis* PCC 6803 transformed with the expression vector pVZ322*mcyF*. The cells were grown under high-intensity light conditions, as described in the Experimental section. RE, crude protein extract; E250, protein eluted with lysis buffer containing 250 mM imidazole. The position of a molecular-mass marker is shown to the left of the gel in (**A**) and (**B**).

downstream of an open reading frame highly similar to that of *N*-acetylmuramoyl-L-alanine amidases, and shows 71% identity and 83% similarity with the glutamate racemase of *Synechocystis* PCC 6803 (Figure 7). We therefore propose that the enzyme encoded by this gene is responsible for the formation of

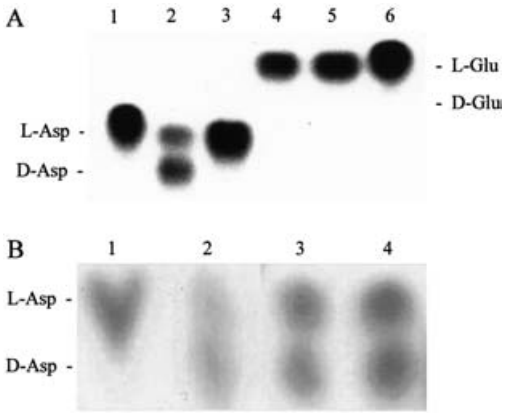

Figure 6 Separation of the enantiomers of aspartate and glutamate on chiral TLC plates

(**A**) An autoradiograph is shown. Cell-free extracts of the *Synechocystis* PCC 6803 wild-type and transconjugant strains were incubated with ^{14}C-labelled aspartate and glutamate, proteins were precipitated and the supernatant was evaporated, as described in the Experimental section. Lane 1, L-aspartate; lane 2, transconjugant with L-aspartate; lane 3, wild-type with L-aspartate; lane 4, wild-type with L-glutamate; lane 5, transconjugant with L-glutamate; lane 6, L-glutamate. The control D-enantiomers were unlabelled, and therefore are not revealed on the autoradiograph; their respective positions, however, are indicated. (**B**) Cell-free extracts of the *Synechocystis* PCC 6803 wild-type and transconjugant strains and purified McyF were incubated as above with unlabeled L-aspartate. Lane 1: wild-type, lane 2: transconjugant, lane 3: purified McyF, lane 4: D- and L-aspartate.

D-glutamate used in peptidoglycan and also in microcystin biosynthesis. The sequence has been deposited in the EMBL database (accession number AJ549183).

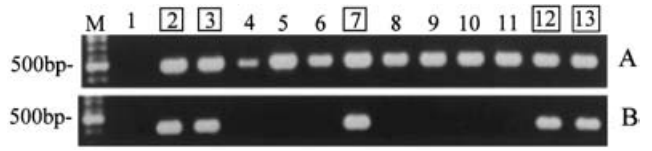

Figure 8 PCR on genomic DNA of 12 *M. aeruginosa* strains using primers corresponding to conserved regions in the cyanobacterial glutamate racemases (A) and *mcyF* (B)

PCR on genomic DNA of 12 *M. aeruginosa* strains was performed as described in the Experimental section using primers corresponding to conserved regions in the cyanobacterial glutamate racemases and *mcyF* (**A** and **B** respectively). Microcystin-producing strains are boxed. M, marker lane; lane 1, control (no template DNA); lane 2, strain 199; lane 3, strain 205; lane 4, strain 265; lane 5, strain 269; lane 6, strain HUB 524; lane 7, strain NIES 102; lane 8, strain NIES 104; lane 9, strain NIES 111; lane 10, strain NIES 44; lane 11, strain PCC 7005; lane 12, strain PCC 7941; lane 13, strain PCC 7806.

To investigate experimentally the hypothesis of an essential function of the glutamate racemase gene, we have used primers on the basis of highly conserved sequence motifs of the L-glutamate racemases from *M. aeruginosa* PCC 7806 and *Synechocystis* PCC 6803 to probe the genomes of several *Microcystis* strains, producers and non-producers of microcystin, for the presence of homologues of the identified L-glutamate racemase gene. Likewise, we have used primers for *mcyF* from *M. aeruginosa* to detect *mcyF* homologues. Remarkably, even though all strains examined carry a copy of the glutamate racemase gene, only those producing microcystin possessed a *mcyF* homologue (Figure 8). Thus the data support our conclusion that the aspartate racemase McyF is an enzyme specifically dedicated to microcystin biosynthesis. The newly identified glutamate racemase has been

```
M.GluR  MRESQLSPIGVFDSGVGGLTV..LRELYKQLP.....QES...ILYFADTARLPYGTRSKGEIIEFVREILTWMGE  66
S.GluR  MREPQRSRIGVFDSGVGGLTV..LRELYRQLP.....KES...ILYFGDTARLPYGKRSPQVILQYVREILTWMAA  66
M.McyF  M.KTKLPILGVLGGMGPVVTAEFLKSIYEYNPFIDKEQESPNVIVFSFPSAPDRTGSIDSGKEREFIDFIQVNLEH  75

M.GluR  .HRVKMVIMACNTSSALALEEVRAEFKDLPILGVILPGAKTAVKQGKRIGVISTPATAKSNAYKQAIQEIDPTAQV 141
S.GluR  .EEVKMVIMACNTSSALALETVQQEF.NMPILGVILPGARAAVRQGRRIGVISTPATAASNAYRHAIHEITPDALV 140
M.McyF  LNKLADCIVIGCCTAHYALPQIPENLKD.KLISLIKIADQELQEYNEPTLLLASTGTYQKKLFQEGCTTADLIISL 150

M.GluR  WQIPCPEFVPLIEANRIFDPYTTQVAREYLQPLLAENIDTLVYGCTHYRHL.SPVLRRLLPSSVRLVDPAAAVVRA 216
S.GluR  WQMACPEFVPLIEQNRLHDPYTLEVAKGYLQPLLDADIDTLVFGCTHYRHL.TPVFQQILPSHIRLVDPASHVVKA 215
M.McyF  SESDQKLIHEMIYKVLKRGHDPLSILRDIEALLEKYNTRSYISGCTEF.HLLTKSLKLKGIDSIKAIDPLSTIAQ. 224

M.GluR  AEKELELLGLKNPETSLPTNFAVSGDPDTFARLSRQWLGFSPTVEKVYLPVPVL..... 270
S.GluR  ARQELEVMGLRNSEMSIATRFTVSGCPQQFAELSHQWLGFTPMVEKISLPCLSSICPQPLEIRE 279
M.McyF  ...NFSQLIIKQAQVDLVTDCHQPSNPKSP..... 251
```

Figure 7 Sequence alignment of McyF and glutamate racemases from *M. aeruginosa* PCC 7806 and *Synechocystis* PCC 6803

Accession numbers are featured in parentheses. M.McyF, *M. aeruginosa* PCC 7806 McyF (Q9RNB4); S.GluR, *Synechocystis* PCC 6803 glutamate racemase (S74824); M.GluR, *M. aeruginosa* PCC 7806 putative glutamate racemase (AJ549183). Conserved residues are shown in grey boxes. All residues strictly conserved in glutamate racemases, as defined by Glavas and Tanner [19], are marked with an asterisk. Ser-79 and Arg-190, which replace alanine and proline in both cyanobacterial glutamate racemases at the respective positions, are shown by '+' signs.

shown to be an enzyme that is essential for all strains, which is in agreement with a role in peptidoglycan biosynthesis. Further studies are needed to explore its suggested additional role in microcystin biosynthesis in hepatotoxic *Microcystis* strains.

We thank Dr Michael Hisbergues and Dr Thomas Hübschmann (Institute for Biology, Humboldt University, Berlin) for gifts of McyF antibodies and vector pPbs respectively. We thank Dr Ralf Dieckmann (Institute for Chemistry, Technical University, Berlin) for useful discussions of the manuscript, and Jana Müller (Institute for Biology, Humboldt University, Berlin) for skilled technical assistance. We also thank S. Ferris, A. M. Castets, C. Pichon, L. Frangeul, A. Marcel, P. Glaser and S. Cole (all at the Génopole Institut Pasteur, Paris), who were involved in the sequencing project. This work was supported by the EU project CYANOMCYCES (QLK3-CT-2000-00131) to T. B. and H. v. D.

REFERENCES

1 Sivonen, K. and Jones, G. (1999) Cyanobacterial toxins. In Toxic Cyanobacteria in Water. A Guide to their Public Health Consequences, Monitoring and Management (Chorus, I. and Bartram, J., eds.), pp. 41–111, World Health Organization, E. Spon and F. N. Spon, London and New York

2 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

3 Ito, E., Kondo, F., Terao, K. and Harada, K.-I. (1997) Neoplastic nodular formation in mouse liver induced by repeated intraperitoneal injections of microcystin-LR. Toxicon **35**,1453–1457

4 Zhou, L., Yu, H. and Chen, K. (2002) Relationship between microcystin in drinking water and colorectal cancer. Biomed. Environ. Sci. **15**, 166–171

5 Tillett, D., Dittmann, E., Erhard, M., von Döhren, H., Börner, T. and Neilan, B. A. (2000) Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide–polyketide synthetase system. Chem. Biol. **7**, 753–764

6 Nishizawa, T., Asayama, M., Fujii, K., Harada, K. and Shirai, M. (2000) Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp. J. Biochem. (Tokyo) **126**, 520–529

7 Nishizawa, T., Ueda, A., Asayama, M., Fujii, K., Harada, K., Ochi, K. and Shirai, M. (2000) Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin. J. Biochem. (Tokyo) **127**, 779–789

8 Nishizawa, T., Asayama, M. and Shirai, M. (2001) Cyclic heptapeptide microcystin biosynthesis requires the glutamate racemase gene. Microbiology **147**, 1235–1241

9 Rippka, R., Desrulles, J., Waterbury, J. B., Herdman, M. and Stanier, R. Y. (1979) Generic assignment, strain histories and properties of pure cultures of cyanobacteria. J. Gen. Microbiol. **11**, 1–61

10 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

11 Franche, X. and Damerval, T. G. (1988) Tests on *nif* probes and DNA hybridization. Methods Enzymol. **167**, 803–808

12 Hübschmann, T., Börner, T., Hartmann, E. and Lamparter, T. (2001) Characterization of the Cph1 holo-phytochrome from *Synechocystis sp*. PCC 6803. Eur. J. Biochem. **268**, 2055–2063

13 Zinchenko, V. V., Piven, I. V., Melnik, V. A. and Shestakov, S. V. (1999) Vectors of the complementation analysis of cyanobacterial mutants. Russ. J. Genet. **35**, 228–232

14 Elhai, J. and Wolk, C. P. (1988) Conjugal transfer of DNA to cyanobacteria. Methods Enzymol. **167**, 747–754

15 Bradford, M. M. (1976) A rapid and sensitive method for the quantification of microgram of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–254

16 Lamparter, T., Podlowski, S., Mittmann, F., Schneider-Poetsch, H., Hartmann, E. and Hughes, J. (1995) Phytochrome from protonemal tissue of the moss *Ceratodon purpureus*. J. Plant Physiol. **147**, 426–434

17 Tanner, M. E., Gallo, K. A. and Knowles, J. R. (1993) Isotope effects and the identification of catalytic residues in the reaction catalysed by glutamate racemase. Biochemistry **32**, 3998–4006

18 Liu, L., Iwata, K., Kita, A., Kawarabayasi, Y., Yohda, M. and Miki, K. (2002) Crystal structure of aspartate racemase from *Pyrococcus horikoshii* OT3 and its implications for molecular mechanism of PLP-independent racemization. J. Mol. Biol. **319**, 479–489

19 Glavas, S. and Tanner, M. E. (2001) Active site residues of glutamate racemase. Biochemistry **40**, 6199–6204

20 Chiaramonte, S., Giacometti, G. M. and Bergantino, E. (1999) Construction and characterization of a functional mutant of *Synechocystis* 6803 harbouring a eukaryotic PSII-H subunit. Eur. J. Biochem. **260**, 833–843

21 Yamauchi, T., Choi, S. Y., Okada, H., Yohda, M., Kumagai, H., Esaki, N. and Soda, K. (1992) Properties of aspartate racemase, a pyridoxal 5′-phosphate-independent amino acid racemase. J. Biol. Chem. **267**, 18361–18364

22 Ashiuchi, M., Tani, K., Soda, K. and Misono, H. (1998) Properties of glutamate racemase from *Bacillus subtilis* IFO 3336 producing poly-γ-glutamate. J. Biochem. (Tokyo) **123**, 1156–1163

23 Michaud, C., Blanot, D., Flouret, B. and van Heijenoort, J. (1987) Partial purification and specificity studies of the D-glutamate-adding and D-alanyl-D-alanine-adding enzymes from *Escherichia coli* K12. Eur. J. Biochem. **166**, 631–637

Received 12 March 2003/23 April 2003; accepted 25 April 2003
Published as BJ Immediate Publication 25 April 2003, DOI 10.1042/BJ20030396