APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Feb. 2002, p. 449–455 Vol. 68, No. 2
0099-2240/02/$04.00+0 DOI: 10.1128/AEM.68.2.449–455.2002

Multiple Alternate Transcripts Direct the Biosynthesis of Microcystin, a Cyanobacterial Nonribosomal Peptide

Melanie Kaebernick,[1] Elke Dittmann,[1,2] Thomas Börner,[2] and Brett A. Neilan[1]*

School of Microbiology and Immunology, University of New South Wales, Sydney 2052, Australia,[1] *and Institute for Biology (Genetics), Humboldt University, Berlin, Germany*[2]

Received 6 July 2001/Accepted 24 September 2001

The *mcyABCDEFGHIJ* gene cluster of *Microcystis aeruginosa* encodes the mixed polyketide synthase/nonribosomal peptide synthetase (microcystin synthetase) which is responsible for biosynthesis of the potent liver toxin microcystin. The sequence and orientation of the *mcy* genes have previously been reported, but no transcriptional analysis had been performed prior to this study. The *mcyABCDEFGHIJ* genes are transcribed as two polycistronic operons, *mcyABC* and *mcyDEFGHIJ*, from a central bidirectional promoter between *mcyA* and *mcyD*. Two transcription start sites were detected for both *mcyA* and *mcyD* when cells were exposed to light intensities of 68 and 16 μmol of photons m^{-2} s^{-1}. The start sites, located 206 and 254 bp upstream of the translational start for *mcyD* under high and low light conditions, respectively, indicate long untranslated leader regions. Putative transcription start sites were also identified for *mcyE*, *mcyF*, *mcyG*, *mcyH*, *mcyI*, and *mcyJ* but not for *mcyB* and *mcyC*. A combination of reverse transcription-PCR and rapid amplification of cDNA ends was employed throughout this work, which may have been one of the first transcriptional analyses of a large nonribosomal polyketide gene cluster.

Cyanobacteria are known to produce a wide range of bioactive compounds, including nonribosomally made peptides, polyketides, alkaloids, and lipopolysaccharides. Some of these compounds are neuro- or hepatotoxins, and their impact on human and animal health has resulted in concern worldwide. Microcystin is a hepatotoxic heptapeptide produced by several species of the genera *Microcystis*, *Anabaena*, *Nostoc*, and *Oscillatoria.* More than 65 structural isoforms of microcystins with various toxicities have been identified having the common structure cyclo(Adda–D-Glu–Mdha–D-Ala–L-X–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 (27, 31).

Microcystin is synthesized nonribosomally via a mixed polyketide synthase/nonribosomal peptide synthetase system called microcystin synthetase (2, 12, 23). Recently, the gene cluster encoding the microcystin synthetase complex has been identified and sequenced (25, 34). This 55-kb gene cluster consists of six open reading frames (ORFs) with a mixed nonribosomal peptide synthetase/polyketide synthase nature (*mcyA* to *mcyE* and *mcyG*) and four smaller ORFs with putative precursor and tailoring functions (*mcyF* and *mcyH* to *mcyJ*). Catalytic domains in *mcyA* to *mcyE* and *mcyG* are responsible for incorporation of the precursors phenylacetate, malonyl coenzyme A, *S*-adenosyl-L-methionine, glutamate, serine, alanine, leucine, D-methyl-isoaspartate, and arginine. The smaller ORFs encode monofunctional proteins which are putatively involved in O-methylation (McyJ), epimerization (McyF), dehydration (McyI), and cellular localization (McyH) (24, 34).

* Corresponding author. Mailing address: School of Microbiology and Immunology, University of New South Wales, Sydney 2052 NSW, Australia. Phone: 61-2-93853235. Fax: 61-2-93851591. E-mail: b.neilan@unsw.edu.au.

The 10 ORFs are bidirectionally transcribed from a central 737-bp locus between *mcyA* and *mcyD* and are arranged in what until now has been classified as two putative operons, *mcyABC* and *mcyDEFGHIJ*. It has previously been reported that the *mcyB* and *mcyD* transcript levels are similar when they are determined under various light conditions (18). Nevertheless, the operon structure, the transcription start sites for *mcyA* and *mcyD*, and the presence and locations of putative promoters for individual genes within the cluster have not been investigated in detail (18, 24). Transcriptional analysis of the *mcy* cluster not only should increase our understanding of microcystin synthetase regulation and toxin biosynthesis but may also provide useful insights into other nonribosomal systems, some of which are involved in antibiotic production. Knowledge of the structure of the *mcy* gene cluster may therefore have important implications for further research to produce new antibiotics by combinational biosynthesis.

Transcriptional analysis of gene clusters has traditionally been carried out by Northern blotting, while transcription start sites are commonly mapped by using radioactive primer extension or S1 nuclease mapping (13). However, it has not been possible to detect the *mcyABCDEFGHIJ* genes on Northern blots, probably because of the very low levels and potential lengths of the transcripts. As a result, reverse transcription (RT)-PCR was used to detect mRNA transcripts in the *mcyABCDEFGHIJ* cluster. This method has previously been used for transcriptional analysis of cyanobacterial hydrogenase genes in *Anabaena* sp. (3, 5) and GTPase/ATPase genes in *Synechocystis* sp. (16).

Transcription initiation points can be mapped by rapid amplification of cDNA ends (RACE), also known as ligation-anchored PCR (30). This method has only recently been modified so that it can be used with prokaryotes (33). It involves production of the desired cDNA by RT, linking of a modified anchor oligonucleotide to the 5′ end of cDNA, and successive

TABLE 1. Oligonucleotides used for RT and PCR

Primer			Sequence (5′-3′)
Gene	Type[a]	Designation	
Primers used for RT and RACE-PCR			
mcyA	R_2	PMR2	CTA GAG TAG TTA CTG GAT CTA
mcyB	R_2	mcyBpR2	AAC CCT TGT TGA CGC TCT GT
mcyC	R_2	Elke10	CTG CCA AGA GTT CAG CAT CT
mcyD	R_2	Topaz4	TTG TTG CCA TTC CTG CAT
mcyE	R_2	mcyEpR2	CGT AAA CCT TCA GGG CAT GA
mcyF	R_2	mcyFpR2	TGC AGC CTT CCT GAA ATA ATT
mcyG	R_2	mcyGpR2	AGC AGT CAT GGA CAG AAT ATT
mcyH	R_2	Ebony5	GAC GCT CAA CAT AAC TGG AA
mcyI	R_2	mcyIpR2	GTT GCA TCT GAT TGC GAA TTA
mcyJ	R_2	mcyJpR2	AAT CCG CTA CAG CCA ACC TT
Primers used for RACE-PCR and RT-PCR			
mcyA	R_1	Promi5R	GGT AAT CTA TTG AAA CCA GAT G
mcyB	R_1	mcyBpR	CGA GTG ACG CTC TAC AAC TT
mcyC	R_1	mcyCpR	CCA GTT CTT GAA AGG CTG CT
mcyD	R_{1a}	MpR60R	TTT GTT GAG ACT CGA AAG CTT TTA
mcyD	R_{1b}	MpR15R	GTT TTC GGA TAA GTT CTT TTT ATC
mcyE	R_1	mcyER	GAC TTC CAC TAT AGC ATC AAG
mcyF	R_1	mcyFR	CTA TCA ATA GAA CCT GTG CGA
mcyG	R_1	mcyGR	ACA AGC CCA TAG TGC GGG AAT
mcyH	R_1	abcpR	TCT CAG AAA CTT AGA GGC ATT
mcyI	R_1	Ebony3	GCC TCA TTA ATT TCC TCC GGT
mcyJ	R_1	mcyJpR	TGA TCT CCG GCT TTA AGT TC
Primers used for RT-PCR			
mcyB	F	Ruby43	CGG TTA TTT TGC GAC TCC
mcyC	F	Elke12	GAA AAA TTC ATC GTC AAG CT
mcyE	F	mcyDF	GTC TGT CTC TAT AGA TGA ACG
mcyF	F	mcyEF	GAT CTG GCC TTG GAC GGA ATT
mcyG	F	mcyFF	TTG ACC ATC GCA CCT CTA TCA
mcyH	F	abcpF	GAC TTT CCT AAT ATT AGA CGC
mcyI	F	mcyIpF	CTC CCT TAG AAC AAT TTC TCT
mcyJ	F	mcyJpF	GAG CAA CTT TGC TCA AAG ATT
Primers used for RACE ligation and RACE-PCR			
		DT89	CGC CAT TTC CAC CTT CTC TTC
		DT88	GAA GAG AAG GTG GAA ATG GCG TTT TGG

[a] R, reverse primer (R_1 is upstream of R_2; R_{1a} and R_{1b} were used interchangeably, leading to the same results); F, forward primer.

amplification with a primer complementary to the anchor and primers specific for the cDNA upstream of the reverse transcriptase primer sequence. The addition of the anchor places a known unique sequence at the unmapped 5′ end of the cDNA, which can be identified by DNA sequencing of the nested fragment.

Using RT-PCR and RACE, we performed a complete transcriptional analysis of the *mcyABCDEFGHIJ* cluster, including

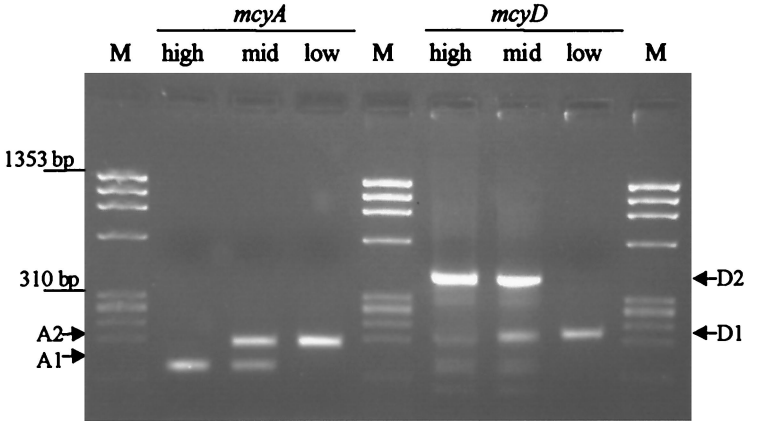

FIG. 1. RACE products identified for *mcyA* and *mcyD* from RNA extracted from cells grown with 31 μmol of photons m^{-2} s^{-1} (intermediate light conditions) and exposed prior to sampling to high light conditions (68 μmol of photons m^{-2} s^{-1}) and low light conditions (16 μmol of photons m^{-2} s^{-1}). ϕX174 DNA/*Hae*III (Promega) was used as a size marker (lanes M). The bands represent 1,353, 1,078, 872, 603, 310, 281, 271, 234, and 194 bp (from top to bottom).

an analysis of putative transcription initiation points for the *mcy* genes and alternate transcription start sites for *mcyA* and *mcyD* under different light conditions. Our results indicate that complex gene regulation occurs in hybrid nonribosomal peptide-polyketide biosynthesis involving both multiple and alternate messages.

MATERIALS AND METHODS

Cyanobacterial strains and culturing. The axenic strain *Microcystis aeruginosa* PCC7806 (Braakman Reservoir, The Netherlands) was provided by the Pasteur Culture Collection, Paris, France (28). Cells were grown in BG11 (Sigma-Aldrich, Steinheim, Switzerland) in glass cylinders (diameter, 5.5 cm) with continuous aeration at 25°C and with continuous light (31 μmol of photons m^{-2} s^{-1}). For the light experiments the cultures were exposed to 68 μmol of photons m^{-2} s^{-1} (high light conditions) or 16 μmol of photons m^{-2} s^{-1} (low light conditions) for 6 h before they were harvested.

Sampling, RNA extraction, and DNase treatment. Both the sampling and RNA extraction procedures used have been described previously (18). After extraction with Trizol (Gibco BRL, Life Technologies, Rockville, Md.), RNA was column purified (SV total-RNA isolation system; Promega, Madison, Wis.). DNA was removed by treatment with DNase (Gibco BRL), followed by phenol-chloroform-isoamyl alcohol (25:24:1) extraction. The amount and purity of the RNA were determined from the optical densities at 260 and 280 nm and by electrophoresis on a 1% formaldehyde gel.

RT. RT for production of cDNA for RACE or RT-PCR analysis (see below) was carried out by using Superscript II (Gibco BRL). RNA (0.5 to 1 μg) was added to 1× Superscript II buffer containing 10 pmol of reverse primer R_2 (Table 1), each deoxynucleoside triphosphate at a concentration of 0.2 mM, 1 mM dithiothreitol, and enough H_2O to bring the final volume to 45 μl. The Superscript II enzyme (200 U diluted in enough H_2O to bring the final volume to 5 μl) was added only after stepwise primer annealing by a touchdown incubation procedure consisting of 70°C for 2 min, followed by 65°C for 1 min, 60°C for 1 min, 55°C for 1 min, 50°C for 1 min, and 45°C for 1 min, using a thermocycler (Perkin-Elmer). After the enzyme was added, the reaction mixture was incubated at 42°C for 30 min, which was followed by five cycles consisting of 50°C for 1 min, 53°C for 1 min, and 56°C for 1 min. The enzyme was omitted from reaction mixtures that were used as negative controls in the RT-PCR analysis. Untranscribed RNA was removed by alkaline cleavage; 1 μl of 0.5 M EDTA and 12.5 μl of 0.2 M NaOH were added directly to the RT reaction mixture, and then the preparation was incubated at 68°C for 5 min. The reaction was stopped by adding 12.5 μl of 1 M Tris-HCl (pH 7.4). cDNA was precipitated with 60 μl of isopropanol, 5 μl of 3 M sodium acetate, and 20 μg of glycogen. The pellet was dried, resuspended in 20 μl of TE (10 mM Tris-HCl [pH 7.4], 1 mM EDTA [pH 8.0]), and stored at −20°C.

RACE. Transcription start sites were identified by ligation-anchored PCR (33). In this procedure, T4 RNA ligase was used to ligate a 5′-phosphorylated, 3′ end cordycepin-blocked anchor oligonucleotide, DT88 (Table 1), to the single-stranded cDNA. The ligation conditions were as follows: 1× RNA ligase buffer (Fermentas), 0.4 μM DT88, 10 μl of cDNA sample, and 1 μl (3 U) of RNA ligase in a 12-μl (final volume) mixture. The reaction mixture was incubated overnight at room temperature. The PCR was performed directly with the DT88-ligated cDNA (1 μl), 10 pmol of DT89, and R_2 primers (Table 1). The PCR was hot started by adding 5 μl of water containing 1 U of *Taq* DNA polymerase (Biotechnology International) after an initial 2-min denaturation step at 90°C. Then the following touchdown program was used: 95°C for 10 s and 70 to 60°C (1 min each) for 10 cycles, followed by 15 cycles of 95°C for 10 s and 60°C for 1 min. A second, nested PCR was carried out under the same conditions with a reverse primer (R_1) internal to the primer originally used for cDNA synthesis (R_2). The resulting RACE products were visualized by agarose gel electrophoresis, gel purified, and sequenced to determine the transcription start sites for each *mcy* gene. All RACE assays were repeated several times with up to six different RNA samples. Analyses of the *mcyA* and *mcyD* transcription start sites were repeated with RNA samples from three separate cultures grown under the conditions described above.

RT-PCR. The presence of transcripts in selected gene regions was checked by RT-PCR. The reverse primer (R_1) (Table 1) used to do this was situated upstream of the original reverse primer used for the RT reaction (R_2). PCRs were performed with an initial denaturation step of 95°C for 2 min, followed by 30 cycles of 94°C for 10 s, 50°C for 20 s, and 72°C for 1 min, using 1 μl of cDNA prepared as described above, 10 pmol of reverse primer (R_1), and 10 pmol of forward primer (Table 1).

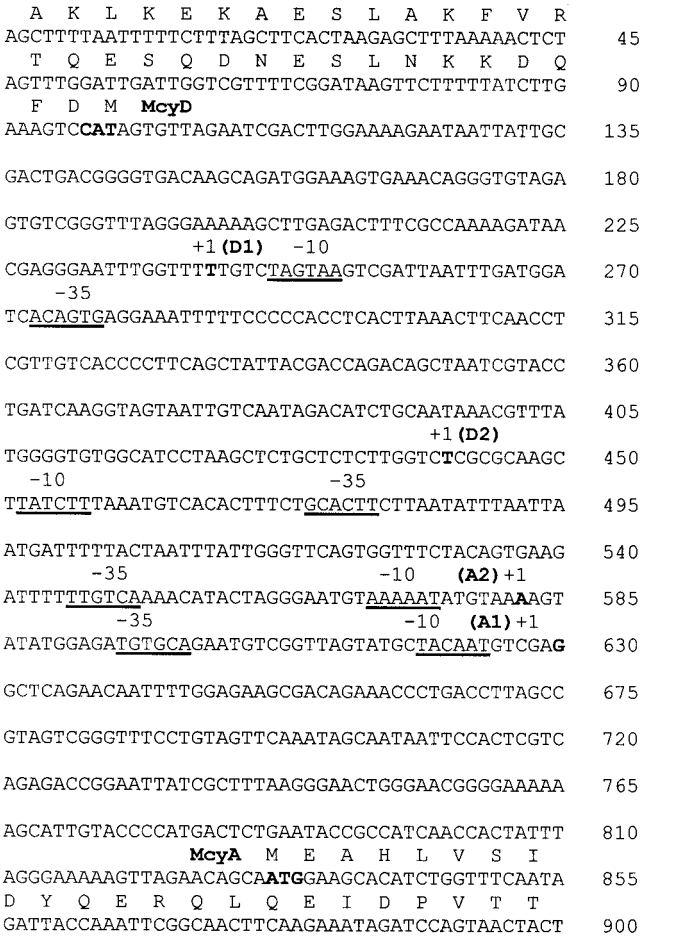

FIG. 2. Nucleotide sequence of the region separating *mcyA* and *mcyD*, including portions of each ORF. The site of transcription initiation for each gene is indicated by +1, and the translation start codon is indicated by boldface type. Putative promoter sequences at −10 and −35 are underlined.

RESULTS

Transcription initiation sites for *mcyA* and *mcyD* are dependent on light. Both *mcyA* and *mcyD* had two transcription start points, which appeared to be dependent on the light conditions under which the harvested culture grew. Analysis of RNA from cultures grown with 31 μmol of photons m^{-2} s^{-1} (intermediate light conditions) showed that both RACE products, corresponding to the transcription start sites found under high and low light conditions (Fig. 1), were present.

For *mcyA*, two transcription start sites were identified, which were 206 bp (A_H) and 254 bp (A_L) upstream of the translation start site under high and low light conditions, respectively. For *mcyD*, low and high light conditions resulted in the use of transcription start sites which were 143 bp (D_L) and 342 bp (D_H) upstream of the initiating methionine (Fig. 2). Promoter motifs at −10 and −35 bp upstream of the transcription initiation site exhibited similarity to the *Escherichia coli* σ^{70} consensus promoter sequences (Fig. 2).

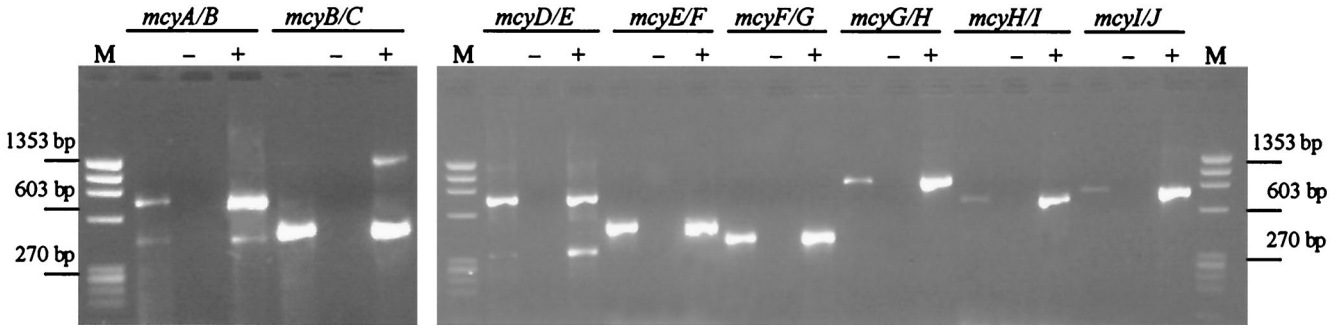

FIG. 3. RT-PCR analysis showing the presence of transcripts between *mcyA* and *mcyB*, between *mcyB* and *mcyC*, between *mcyD* and *mcyE*, between *mcyE* and *mcyF*, between *mcyF* and *mcyG*, between *mcyG* and *mcyH*, between *mcyH* and *mcyI*, and between *mcyI* and *mcyJ*. For each gene, negative controls in which RNA was used for the PCR (lanes −) and positive controls in which DNA was used for the PCR (lanes +) were included. ϕX174 DNA/*Hae*III (Promega) was used as a size marker (lanes M). The bands represent 1,353, 1,078, 872, 603, 310, 281, 271, 234, and 194 bp (from top to bottom).

***mcyABCDEFGHIJ* gene cluster is arranged as two polycistronic operons, *mcyABC* and *mcyDEFGHIJ*.** RT-PCR analysis revealed the presence of transcripts in the noncoding regions between *mcyA* and *B*, *mcyB* and *C*, *mcyD* and *E*, *mcyE* and *F*, *mcyF* and *G*, *mcyG* and *H*, *mcyH* and *I*, and *mcyI* and *J* (Fig. 3). The amplicons obtained from the cDNA were the same sizes as the amplicons obtained from the genomic DNA. Negative control experiments, which revealed a complete absence of DNA in the RNA samples, were performed by subjecting a duplicate RNA sample to the same experimental procedures but omitting the RT enzyme during RT.

Detection of putative intercistronic promoters. The RACE assay was also employed to check for potential individual ORF promoters internal to the polycistronic transcripts. RACE products were identified for most of the microcystin synthetase genes. Some of these products were repeatedly amplified to obtain a quantity useful for sequencing and visualization by gel electrophoresis (Fig. 4). The single amplicons observed for *mcyF*, *mcyG*, *mcyH*, *mcyI*, and *mcyJ* indicated that there were additional intercistronic transcription start sites for these genes (Fig. 4 and 5). Putative transcription start sites and promoters for each of the genes are shown in Table 2. Interestingly, four different RACE products were found for *mcyE*. These products were identified repeatedly with the same and successive RNA samples. All four start sites exhibited low levels of similarity to the *E. coli* σ^{70} promoter motifs (Table 2). A fifth *mcyE* band below band 2 (Fig. 4) was too small to extend upstream of the translation start site of *mcyE*. This band was also checked by sequencing and may have indicated that incomplete cDNA synthesis, RNase processing, or transcript degradation occurred. There were no RACE products identified for *mcyB* and *mcyC*.

DISCUSSION

Findings presented here showed that the microcystin synthetase gene cluster, *mcyABCDEFGHIJ*, is transcribed as two polycistronic transcripts, *mcyABC* and *mcyDEFGHIJ*, from a bidirectional central promoter situated between *mcyA* and *mcyD*. For both *mcyA* and *mcyD*, two transcription start sites were identified, and the use of these start sites appeared to depend on the light intensity. Furthermore, most of the polyketide and tailoring *mcy* genes (*mcyE*, *mcyF*, *mcyG*, *mcyH*, *mcyI*, and *mcyJ*) may also possess individual promoters (Fig. 5).

Light-dependent transcription start sites were identified for *mcyA* and *mcyD* when cells were exposed to 68 μmol of photons m^{-2} s^{-1} (high light conditions) and 16 μmol of photons m^{-2} s^{-1} (low light conditions) prior to harvesting and RNA extraction. When cells were grown with constant light of 31 μmol of photons m^{-2} s^{-1}, both transcription start sites were detected (Fig. 1). The use of multiple promoters for expression of the same cyanobacterial gene has previously been reported for *glnA* (glutamine synthetase) and *argD* (*N*-acetylornithine aminotransferase) in *Anabaena* sp. strain PCC7120 (15, 35),

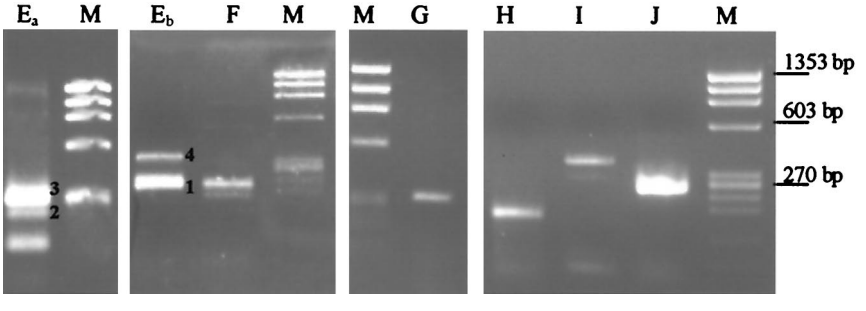

FIG. 4. RACE products detected for *mcy* genes in the polycistronic *mcyDEFGHIJ* transcript (*mcyE*$_a$, *mcyE*$_b$, *mcyF*, *mcyG*, *mcyH*, *mcyI*, and *mcyJ*). Lanes E_a and E_b show results obtained with two different RNA samples. *mcyE* RACE products 1, 2, 3, and 4 were purified and sequenced (Table 2). ϕX174 DNA/*Hae*III (Promega) was used as a size marker (lanes M). The bands represent 1,353, 1,078, 872, 603, 310, 281, 271, 234, and 194 bp (from top to bottom).

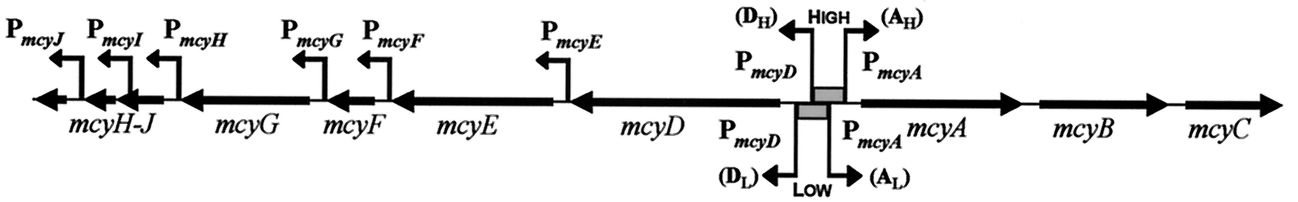

FIG. 5. Organization of the microcystin synthetase gene cluster, *mcyABCDEFGHIJ*, showing putative promoters for *mcyEFGHIJ* (P_{mcyE} to P_{mcyJ}) and the alternate promoters identified for *mcyA* and *mcyD* (P_{mcyA} and P_{mcyD}) under high light conditions (A_H and D_H) and low light conditions (A_L and D_L).

for *zwf* (glucose-6-phosphate dehydrogenase) in *Nostoc punctiforme* (32), and for the *petH* gene encoding ferredoxin: $NADP^+$ reductase in a heterocyst-forming *Anabaena* sp. (36). In the case of *petH*, one of the promoters directs constitutive transcription under different nitrogen nutrition conditions, while the other is induced in cells subjected to nitrogen stepdown and in nitrogen-fixing filaments (36). Similarly, *Microcystis* cells may use different promoters depending on exposure of the cells to different light conditions. Light-dependent transcription has previously been demonstrated for *mcyB* and *mcyD* and may indicate that microcystin has a function that is related to irradiance (18).

One of the initial steps in which transcription initiation is regulated is the interaction of the RNA polymerase with conserved promoter sequences situated −10 and −35 bp upstream of the transcription initiation site. Most cyanobacterial −10 promoter sequences are highly conserved with the consensus *E. coli* σ^{70} promoter motif (TANNNT) (11). The −35 motif, however, shows only weak conservation with the *E. coli* σ^{70} consensus sequence (TTGACA) in approximately one-half of the promoters mapped for different cyanobacterial genera (11). The transcription initiation sites identified under high light conditions for *mcyA* and *mcyD* indicated that there were conserved −10 motifs consisting of TACAAT and TCTATT, respectively (Fig. 2). Conservation was less prevalent in the −10 basal promoter sequences of *mcyA* (AAAAAT) and *mcyD* (CTGAAT) transcripts generated under low light conditions. Less conserved promoter motifs generally indicate weaker transcription. However, the −35 regions of the low-light-conditions transcripts of *mcyA* (TTGTCA) and *mcyD* (GTGACA) exhibited higher levels of similarity to the *E. coli* σ^{70} consensus than those identified for the *mcyA* (TGTGCA) and *mcyD* (TTCACG) transcripts under high light conditions (Fig. 2).

The less conserved −35 motif in the *mcyA* and *mcyD* high-light-conditions transcripts could indicate involvement of *cis* factors superimposed on the promoter region, as found in regulated *E. coli* σ^{70} promoters (10). In these types of promoters, the binding sites of activators are close to the −35 region, leading to less conservation of this motif. Thus, it may be speculated that *mcyABC* and *mcyDEFGHIJ* are constitutively transcribed from the weak promoters exhibited under low light conditions and that an activator is involved in transcription initiation under high light conditions. Initiation from the alternate promoters under high light conditions may lead to increased transcription under these conditions, as previously determined for *mcyB* and *mcyD* (18).

RACE products identified for *mcyEFGHIJ* indicated that intercistronic promoters were present. Promoter elements at −10 and −35, however, exhibited only weak similarity to the consensus *E. coli* σ^{70} (Table 2). The absence of typical promoter elements could indicate false predictions of transcription start sites due to incomplete processing of a larger transcript during RT or due to RNA cleavage (6, 14). This may be true especially for the four transcription start sites detected for *mcyE*. However, weak conservation of the −10 and −35 motifs or the complete absence of −35 regions has previously been documented in genes from various cyanobacterial genera, including the *glnA*(*IV*) and *psbB*(*I*)(*II*) genes from *Anabaena* sp. strain PCC7120, in which the transcription initiation site was verified by in vitro transcription assays (19, 35).

Internal promoters have been identified in the polycistronic operons of several other bacterial gene clusters (9, 22, 38). The

TABLE 2. Putative intercistronic transcription start points, +1 nucleotides, promoter sequences at −10 and −35, and locations of the transcription start points and the translation start sites (ORF) in the *mcyABCDEFGHIJ* cluster (1 to 55,000 bp)

Gene	Location of:		Sequence		
	Transcription start point (bp)[a]	ORF (bp)[a]	−35	−10	+1
mcyE$_1$	24175	24173	ATTTATGTTTAACGGCGTATAAATTTAGAAATTATAGAGAATAGGG		
mcyE$_2$	24187	24173	CACAGCATCCTAATTTATGTTTAACGGCGTATAAATTTAGAAATTA		
mcyE$_3$	24232	24173	GCTTAACTCTCTGGGGTTAAATTTGGGTTTATATTTACTGATAGCA		
mcyE$_4$	24288	24173	AAAATTATTTTTCAAAAGCCATGAGATGTGGTAGGATAGGTAGATG		
mcyF	13703	13678	GAGGGGCAGGATATGTTCACAGAAGATTTTAATTTTTAGTCAACAG		
mcyG	12883	12791	CCTTGACCATCGCACCTCTATCAGGTACAATTCAAACTTGAGTTGA		
mcyH	4692	4672	CTCCTTACTGGTATCCAACTGATCCTAAAGGAAGGACATTTTCAGA		
mcyI	3167	3017	CTTATATCAGCAGTTACAAAAAGCACAGACAACTTTTATTAGCGTA		
mcyJ	1909	1828	CTGCAAAATCCTTAGGAGGATAATAATTTGGAAAACATCGGCAAAA		
E. coli σ^{70}			TTGACA	TATAAT	

[a] Location starting from the 3′ *mcyJ* end (accession no. AF183408).

his operon in *E. coli* and *Salmonella enterica* serovar Typhimurium, for example, consisting of eight polycistronic genes, contains two internal promoters within *hisC* and *hisF*, as well as the primary promoter at the 5′ end (9). The distal cistrons of the operon are expressed from the two internal promoters, in addition to the primary element. Expression from these alternate promoters appears to depend on various physiological states (17). Similarly, it could be speculated that the intercistronic *mcy* promoters are functional or more active only under particular growth conditions.

It has also been proposed that internal promoters may have a function, especially in long and generally unstable transcripts, in ensuring adequate expression of the distal genes of the operon or cluster (26). This has been suggested for the 14-kb *cap1* gene cluster of *Staphylococcus aureus,* which is transcribed from one strong primary promoter and five weak internal promoters which exhibit significant activity when the primary promoter is removed (26). The internal promoter of a particular gene, such as *ilvEp* in *E. coli*, can contribute as much as 90% of the total expression (21). However, as transcription through a promoter inhibits its activity (promoter occlusion), the internal promoters probably function only when the primary transcript is interrupted, degraded, or otherwise repressed (1). This may also be true for the large 55-kb *mcyABCDEFGHIJ* gene cluster. RACE products of the *mcyEFGHIJ* genes were identified in only approximately one-half of the RNA samples tested. Under conditions which favor transcription from the primary promoters at *mcyA* and *mcyD* and high transcript stability, the internal promoters may be subject to promoter occlusion. Thus, without knowing the precise conditions which lead to activity of the internal promoters rather than transcription from the primary promoter, we could not consistently reisolate the RACE products which identify the *mcy* internal promoters.

Alternatively, our results may be explained by the presence of processing (cleavage) sites in the large transcripts. Processing could produce monocistronic mRNAs for translation, as observed in chloroplasts (4, 29). The internal promoters identified for the *mcyEFGHIJ* genes are somewhat analogous to the chloroplast gene *petD*. This gene possesses its own promoter for production of a monocistronic transcript. However, it may also utilize the promoter of the upstream gene *petA* and be transcribed as part of a polycistronic message (29). Here, promoter usage depends on RNA processing events and RNA secondary structures at the 5′ end of the gene (4).

Another structural feature of the *mcy* mRNA was the long untranslated leader regions identified for both *mcyA* and *mcyD* (Fig. 2). Transcription initiation sites under high and low light conditions for *mcyA* were located 206 and 254 bp upstream of the ORF start codon, respectively. This provided a longer *mcyABC* transcript when cells were cultured under low light conditions. For *mcyD*, the longer *mcyDEFGHJ* transcript was identified under high light conditions, under which the transcriptional start site was situated 342 bp upstream of the ORF, compared to 143 bp upstream of the ORF when the organism was grown under low light conditions. Light-stimulated differences in the lengths of transcripts have previously been documented in the pea chloroplast (37). Long untranslated leader regions between the basal promoter motifs and the ORF start codon are also characteristic of light-responsive expression of the *psbA* and *psbD* genes in cyanobacteria (7). In these systems, the consensus basal promoter elements are not affected by changes in light intensity, and light-responsive expression is driven by specific *cis* elements in the untranslated leader regions of the mRNA (8). Light-responsive elements have been identified 11 to 84 bp downstream of the transcription start site for *psbDII* and 1 to 41 and 39 downstream of the transcription start site for *psbAII* and *psbAIII*, respectively (7, 20). Such regions are being investigated further for the *mcy* cluster by using a reporter vector system. For the *psb* genes, these regions alone were able to confer a reproducible level of light-responsive expression on a heterologous promoter (8, 20).

In conclusion, we found that the microcystin synthetase gene cluster (*mcyABCDEFGHIJ)* consists of two polycistronic transcripts, *mcyABC* and *mcyDEFGHIJ*, which are transcribed from a central promoter between *mcyA* and *mcyD*. Most interestingly, both polycistronic transcripts have alternate transcription start sites which appear to be light dependent. Furthermore, identification of putative intercistronic transcription start sites for individual *mcy* genes indicated that there is complex gene regulation involving multiple and alternate messages. These findings may prove to be highly significant for our understanding of regulation of the microcystin synthetase gene cluster under particular light conditions and may increase our knowledge concerning a function of microcystin under these growth conditions.

ACKNOWLEDGMENTS

This work was supported by grants from the Australian Research Council and the CRC for Water Quality and Treatment to B.A.N. and M.K., by a grant from the DFG to T.B., and by a grant from the Alexander von Humboldt foundation to E.D.

REFERENCES

1. **Adhya, S., and M. Gottesman.** 1982. Promoter occlusion: transcription through a promoter may inhibit its activity. Cell **29:**939–944.
2. **Arment, A. R., and W. W. Carmichael.** 1996. Evidence that microcystin is a thiotemplate product. J. Phycol. **32:**591–597.
3. **Axelsson, R., F. Oxelfelt, and P. Lindblad.** 1999. Transcriptional regulation of *Nostoc* uptake hydrogenase. FEMS Microbiol. Lett. **170:**77–81.
4. **Barkan, A., M. Walker, M. Nolasco, and D. Johnson.** 1994. A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J. **13:**3170–3181.
5. **Boison, G., H. Bothe, and O. Schmitz.** 2000. Transcriptional analysis of hydrogenase genes in the cyanobacteria *Anacystis nidulans* and *Anabaena variabilis* monitored by RT-PCR. Curr. Microbiol. **40:**315–321.
6. **Boudreau, E., Y. Takahashi, C. Lemieux, M. Turmel, and J.-D. Rochaix.** 1997. The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. EMBO J. **16:**6095–6104.
7. **Bustos, S. A., and S. S. Golden.** 1991. Expression of the *psbDII* gene in *Synechococcus* sp. strain PCC 7942 requires sequences downstream of the transcription start site. J. Bacteriol. **173:**7525–7533.
8. **Bustos, S. A., and S. S. Golden.** 1992. Light-regulated expression of the *psbD* gene family in *Synechococcus* sp. strain PCC7942: evidence for the role of duplicated *psbD* genes in cyanobacteria. Mol. Gen. Genet. **232:**221–230.
9. **Carlomagno, M. S., L. Chiariotti, P. Alifano, A. G. Nappo, and C. B. Bruni.** 1988. Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. J. Mol. Biol. **203:**585–606.
10. **Collado-Vides, J., B. Magasanik, and J. D. Gralla.** 1991. Control site location and transcriptional regulation in *Escherichia coli*. Microbiol. Rev. **55:** 371–394.
11. **Curtis, S. E., and J. A. Martin.** 1994. The transcription apparatus and the regulation of transcription initiation, p. 613–639. *In* D. A. Bryant (ed.), The molecular biology of cyanobacteria. Kluwer Academic Press, Dordrecht, The Netherlands.
12. **Dittmann, E., B. A. Neilan, M. Erhard, H. von Doehren, and T. Boerner.** 1997. Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa*

PCC 7806. Mol. Microbiol. **26:**779–787.
13. **Domanski, T. L., B. L. M. de Jonge, and K. W. Bayles.** 1997. Transcriptional analysis of the *Staphylococcus aureus* gene encoding penicillin-binding protein 4. J. Bacteriol. **179:**2651–2657.
14. **Dominic, B., Y.-B. Chen, and J. P. Zehr.** 1998. Cloning and transcriptional analysis of the *nifUHDK*genes of *Trichodesmium* sp. IMS101 reveals stable *nifD, nifDK* and *nifK* transcripts. Microbiology **144:**3359–3368.
15. **Floriano, B., A. Herrero, and E. Flores.** 1994. Analysis of expression of the *argC* and *argD* genes in the cyanobacterium *Anabaena* sp. strain PCC7120. J. Bacteriol. **176:**6397–6401.
16. **Geisler, M., B. Jakobs, J. Richter, and J. Schumann.** 1996. Cotranscription of a GTPase gene from the cyanobacterium *Synechocystis* PCC6803 and a P-type Ca(2+)ATPase gene. Biochim. Biophys. Acta **1309:**189–193.
17. **Grisolia, V., A. Riccio, and C. B. Bruni.** 1983. Structure and function of the internal promoter (*hisBp*)of the *Escherichia coli* K-12 histidine operon. J. Bacteriol. **155:**1288–1296.
18. **Kaebernick, M., B. A. Neilan, T. Börner, and E. Dittmann.** 2000. Light and the transcriptional response of the microcystin biosynthetic gene cluster. Appl. Environ. Microbiol. **66:**3387–3392.
19. **Lang, J. D., and R. Haselkorn.** 1989. Isolation, sequence and transcription of the gene encoding the photosystem II chlorophyll-binding protein, CP-47, in the cyanobacterium *Anabaena* PCC7120. Plant Mol. Biol. **13:**441–457.
20. **Li, R., and S. S. Golden.** 1993. Enhancer activity of light-responsive regulatory elements in the untranslated leader regions of cyanobacterial *psbA* genes. Proc. Natl. Acad. Sci. USA **90:**11678–11682.
21. **Lopes, J. M., N. Soliman, P. K. Smith, and R. P. Lawther.** 1989. Transcriptional polarity enhances the contribution of the internal promoter, *ilvEp,* in the expression of the *ilvGMEDA* operon in wild-type *Escherichia coli* K12. Mol. Microbiol. **3:**1039–1051.
22. **Lortie, L.-A., M. Pelletier, C. Vadeboncoeur, and M. Frenette.** 2000. The gene encoding IIABLMan in *Streptococcus salivarius* is part of a tetracistronic operon encoding a phosphoenolpyruvate: mannose/glucose phosphotransferase system. Microbiology **146:**677–685.
23. **Moore, R. E., J. L. Chen, B. S. Moore, G. M. L. Patterson, and W. W. Carmichael.** 1991. Biosynthesis of microcystin-LR. Origin of the carbons in the Adda and Masp units. J. Am. Chem. Soc. **113:**5083–5084.
24. **Nishizawa, T., M. Asayama, K. Fujii, K. Harada, and M. Shirai.** 1999. Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp. J. Biochem. (Tokyo) **126:**520–526.
25. **Nishizawa, T., A. Ueda, M. Asayama, K. Fujii, K. Harada, K. Ochi, and M. Shirai.** 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.
26. **Ouyang, S., and C. Y. Lee.** 1997. Transcriptional analysis of type 1 capsule genes in *Staphylococcus aureus*. Mol. Microbiol. **23:**473–482.
27. **Rinehart, K. L., M. Namikoshi, and B. W. Choi.** 1994. Structure and biosynthesis of toxins from blue-green algae (cyanobacteria). J. Appl. Phycol. **6:**159–176.
28. **Rippka, R., and M. Herdman.** 1992. Pasteur culture collection (PCC) of cyanobacterial strains in axenic culture, vol. 1. Institut Pasteur, Paris, France.
29. **Sakamoto, W., N. R. Sturm, K. L. Kindle, and D. B. Stern.** 1994. *petD* mRNA maturation in *Chlamydomonas reinhardtii* chloroplast: role of 5′ endonucleolytic processing. Mol. Cell. Biol. **14:**6180–6186.
30. **Schaefer, B. C.** 1995. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. Anal. Biochem. **20:**255–273.
31. **Sivonen, K.** 1996. Cyanobacterial toxins and toxin production. Phycologia **35:**12–24.
32. **Summers, M. L., and J. C. Meeks.** 1996. Transcriptional regulation of *zwf,* encoding glucose-6-phosphate dehydrogenase, from the cyanobacterium *Nostoc*. Mol. Microbiol. **22:**473–480.
33. **Tillett, D., B. P. Burns, and B. A. Neilan.** 2000. Optimized rapid amplification of cDNA ends (RACE) for mapping bacterial mRNA transcripts. BioTechniques **28:**448, 450, 452–453, 456.
34. **Tillett, D., E. Dittmann, M. Erhard, H. von Doehren, T. Boerner, and B. A. Neilan.** 2000. Structural organisation of microcystin biosynthesis in *M. aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. Chem. Biol. **7:**753–764.
35. **Tumer, N. E., S. J. Robinson, and R. Haselkorn.** 1983. Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. Nature **306:**337–342.
36. **Valladares, A., A. M. Muro-Pastor, M. F. Fillat, A. Herrero, and E. Flores.** 1999. Constitutive and nitrogen-regulated promoters of the *petH* gene encoding ferredoxin:$NADP^+$ reductase in the heterocyst-forming cyanobacterium *Anabaena* sp. FEBS Lett. **449:**159–164.
37. **Woodbury, N. W., L. L. Roberts, J. D. Palmer, and W. F. Thompson.** 1988. A transcription map of the pea chloroplast genome. Curr. Genet. **14:**75–89.
38. **Yanofsky, C., T. Platt, I. P. Crawford, B. P. Nichols, G. E. Christie, H. Horowitz, M. VanCleemput, and A. M. Wu.** 1981. The complete nucleotide sequence of the tryptophan operon of *Escherichia coli*. Nucleic Acids Res. **9:**6647–6668.