Journal of Applied Phycology (2005) 17: 181–185
DOI: 10.1007/s10811-005-6418-2

Cloning and characterization of c-phycocyanin operon from the cyanobacterium *Arthrospira platensis* FACHB341

Jinjie Liu, Xuecheng Zhang*, Zhenghong Sui, Xiaohui Zhang & Yunxiang Mao
College of Marine Life Sciences, Ocean University of China, Qingdao, 266003, P. R. China

*Author for correspondence: e-mail: xczhang@ouc.edu.cn

Received 22 November 2003; revised 2 November 2004; accepted 5 November 2004

Key words: *Arthrospira*, cpc operon, *in vitro* PCR, promoter sequence, *Spirulina*

Abstract

By using *in vitro* PCR method, C-phycocyanin operon of *Arthrospira platensis* FACHB341 was cloned and characterized. The operon consists of 427 bp ussB, 519 bp *cpcB* gene, 111 bp igsB-A region, 489 bp *cpcA* gene, 184 bp ussH region and 357 bp *cpcH* gene. Promoter prediction and signal scan show that there are putative promoter sequences and regulatory elements in ussB and ussH sequences.

Introduction

Arthrospira (previously known as *Spirulina*) is a photosynthetic prokaryotic cyanobacterium, and C-phycocyanin is one of the major photosynthetic biliproteins of it. C-phycocyanin is one of the important components in the electron transfer of photosynthesis. C-phycocyanin also has some functions that are good for human health, such as antioxidant, radical scavenging, anti-inflammatory and anti-cancer properties (Reddy, 2003; Guan & Guo, 2002; Pinero Estrada, 2001; Bhat & Madyastha, 2001), so in recent years it has been drawn more and more attention.

The C-phycocyanin gene sequence was first reported by Pilot & Fox in 1984, in which the oligonucleotide synthesis of alpha and beta subunits of cpc gene of a freshwater cyanobacterum *Agmenellum quadruplicatum* was described. In 1996, Jeamton reported the *cpcA* and *cpcB* gene sequence in *Spirulina platensis*. Yu (2002) and Manen & Falquet (2002) revealed the cpc gene sequence of some *Spirulina* strains and used it as a site to their analyze phylogenetic relationships. There is no more detailed information about the regulatory sequence and cpc operon.

This paper reports the sequence and characterization of cpc operon of *Arthrospira platensis* FACHB341, including *cpcB* gene, *cpcA* gene, *cpcH* gene, and the upstream sequence of *cpcB* gene (ussB) and *cpcH* gene (ussH), the spacers between *cpcB* gene and *cpcA* gene (igsB-A). The aim is to provide more information on the cpc operon of the economically important cyanobacterium, especially on the regulation domain of structure genes, which may offer potential usage in transgenic manipulation system, improving its growth rate and increasing production in mass cultivation of the cyanobacterium.

Materials and methods

Cultivation

Arthrospira platensis FACHB341 was kept in our laboratory. It was cultured in Zarrouk medium at 25 °C. The light intensity was 40 μmol m^{-2} s^{-1}, with 12 light/12 dark cycle. The filaments were harvested at late-log-phase (OD_{560} = 0.8–1.0).

DNA extraction, restriction analysis and cassette ligation

The total DNA was prepared by using the phenol/chloroform extraction protocol (Sambrook, 1989). Five different restriction enzymes, EcoRI, PstI, SalI,

XbaI and Sau3AI, were used for DNA digestion. The digested fragment was purified and then ligated to cassettes according to the manufacture's protocol of LA PCR™ *in vitro* Cloning Kit (TaKaRa). The reaction mixture included 5 μL of digested product, 2.5 μL of cassette, 15 μL of ligation solution I and 7.5 μL of ligation solution Π. Then the mixture was ligated at 16 °C for 1 h and precipitated by 2.5 (v/v) ethanol overnight. After centrifugation, the products were suspended in 5 μL ddH_2O.

Cloning of partial region of cpcB, cpcA gene and IGS region

Partial sequences of *cpcB*, *cpcA* gene and intergenic spacer between *cpcB* and *cpcA* (igsB-A) were amplified from a DNA sample with degenerate primers Y1 and Y2 (Table 1), which were designed according to the gene sequences of *Arthrospira* retrieved from GenBank. PCR amplification was carried out with a Thermal Cycler (FeRoTec Life Express) by using amplicycle reagents in the presence of 50 ng of template DNA, 2.5 mM $MgCl_2$, 0.5 mM dNTP, 0.5 μM of each primer, 2.5 U TaqE and 5 μL of 10 × PCR buffer. Reaction conditions were one initial cycle of denaturation at 94 °C for 8 min, followed by 30 cycles of denaturation at 94 °C for 40 s, annealing at 58 °C for 40 s, and extension at 72 °C for 80 s. After eletrophoresis in 1.0 % low-melting agarose gel in TAE buffer (40 mM Tris-acetate, 1 mM EDTA), PCR products were visualized under ultraviolet light. The resulting targeted bands DNA was recovered according to Sambrook (1989), and the recovery products were then resuspended in 5 μL ddH_2O.

Table 1. DNA sequence of primers used throughout the procedures.

Primer	Sequence
Y1	TTGATGCCTTCACTAAGGTGG
Y2	GTAGTAGCC(T/G)AT(G/A)TCACGAGC
C1	GTACATATTGTCGTTAGAACGCGTAATACGACTCA
C2	CGTTAGAACGCGTAATACGACTCACTATAGGGAGA
S1	TGCATCGCCAGCAAACACAGC
S2	GAGCTGTACTCAGCATTTCGCC
A1	GTCTGATCAGTGGTGCTGCCCAAGCAGTGT
A2	GCAGGGACCTAACTACGCGGCAGACCAACG
PS1	TCCATCGGTGTCACCTCCTAAG
PS2	TGCCCCGTGGTTCATTTCTTCAG

Cloning of upstream sequence of cpcB gene and downstream sequence of cpcA gene

The upstream sequence of *cpcB* gene and downstream sequence of *cpcA* gene was amplified by *in vitro* PCR cloning method. The upstream sequence of *cpcB* gene was obtained by using the chromosome walking in two steps, each of the step including two PCR reactions. For the first step, the primers S1 and C1, followed by the primers S2 and C2 (Table 1) were used in the first and second reaction, respectively. For the second step, primers PS1, PS2, C1 and C2 were used (Table 1) in the two conjoint reactions. Primers S1 and S2 were designed according to the sequence of *cpcB* gene, PS1 and PS2 were designed based on the sequence of the fragment obtained in the former reaction, primer C1 and C2 was corresponding to the cassette in LA PCR™ *in vitro* Cloning Kit.

The PCR reaction was carried out as follows: In firstly reaction, 1 μL of ligation product was put into the reaction reagents, then started with the one initial cycle of denaturation at 94 °C for 10 min, followed by 30 cycles of denaturation at 94 °C for 60 s, annealing at 55 °C for 120 s and extension at 72 °C for 120 s, and in second reaction, PCR was processed with 50 ng of the first PCR product DNA, and the other reagents were the same as above. The targeted bands DNA was recovered, and the recovery products were then resuspended in 5 μL ddH_2O.

The downstream region of the *cpcA* gene was cloned by the same method, with the first primer pair A1 and C1, and the second A2 and C2 (Table 1). Primers A1 and A2 are designed according to the sequence of *cpcA* gene obtained above.

Purification, cloning, sequencing and analysis of PCR target bands

Recovered PCR products were cloned into pMD18-T vector according to the manufacturer's instructions of DNA Ligation Kit Version 2.0 (TaKaRa). Cloned amplicons were screened by M13 universal primers. Sequencing was performed from single-stranded templates by the Sanger dideoxy method. The software of DNASIS (ver 5.0) (Hitachi Software Engineering Co. Ltd.) was used for the DNA analysis. Sequence similarity was searched on databases of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov) (Altschul, 1990). The putative promoter sequences of *cpcB* were predicted in the prokaryotic part of database

Table 2. Putative promoter sequence of *cpcB* gene.

Start	End	Promoter sequence
175	220	ATGAAGTTTGATTAACATTTGTATCAAAATATAAAATTCTTCTCATAAAC
233	278	ATCTTTTAAGATTTCGGAAAGTGTTCTAGGATACTGAAGAAATGAACCAC
334	379	GGAGGTGACACCGATGGATTGATTGTCGTGATCATTCATGGTGTGTCCAA
350	395	GATTGATTGTCGTGATCATTCATGGTGTGTCCAATCCCAACTCAACTCTA

Note: The transcription starts are highlighted. Putative −10 and −35 regions are underlined.

(http://www.fruitfly.com). Potential sites for the binding of transcription regulatory factors were determined by computational analysis of the DNA sequence by using the signal scan (Higo, 1999).

Results

Digestion of total DNA

Among the five enzymes used in this study, Sau3AI can cut the total DNA of *A. platensis* FACHB341, and EcoRI can only partially digest the DNA of the species. The digested products of EcoRI and Sau3AI were recovered and used as templates.

The cpc operon

By integrating of three pieces of DNA fragments, the cpc operon with a length of 2086 bp was obtained (GenBank accession number: AY804216), and the configuration of this operon was illustrated in Figure 1. In the order from 5′ to 3′, the operon consists of 6 parts: the 427 bp of upstream sequence (ussB), the *cpcB* gene (519 bp), which encodes the β subunit of c-phycocyanin, 111 bp intergenic spacer sequence between *cpcB* and *cpcA* (igsB-A), the *cpcA* gene (489 bp), which encodes the α subunit, 184 bp upstream sequence of *cpcH* (ussH), and the partial sequence of *cpcH* gene (357 bp), which related to the synthesis of one of the rod linker polypeptide.

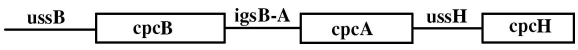

Figure 1. Structure of the cpc operon of *A. platensis* FACHB341. ussB: upstream sequence of *cpcB* gene (1–427 bp); *cpcB*: gene coding for β subunit of C-phycocyanin (428–946 bp); igsB-A: phycocyanin (*cpcB* and *cpcA*) intergenetic spacer region (947–1057 bp); *cpcA*: gene coding for α subunit of C-phycocyanin (1058–1546 bp); ussH: upstream sequence of *cpcH* gene (1547–1729 bp); *cpcH*: gene coding for the linker protein of C-phycocyanin(1730–2086 bp).

The promoter scan in the ussB region indicates that the sequence contains four putative promoter sequences with the same length of 50 bp (Table 2). There are 4 promoter elements TATAAA, GATACTG, GATCATT and TCCAAT in −10 region, and the last one GATACTG is the same as in the promoter P_L sequence of λ bacteriophage. It also has 4 elements TTGATT, TTAAGA, GTGACA and TTGTCG in −35 region. Sequence analysis shows a quite high AT content (63.5%) in the upstream sequence. Repeat sequences and inverted sequences are showed in the GenBank sequence analysis.

Some parts of the promoter region have been shown to correspond to specific transcription factor binding sites. The CAAT element at sites 2, 287, 381 believed to be the CAAT promoter consensus sequence. The TATA element locates in the 5′ upstream sites 39, 70 and 83. The GATA element, a family of nuclear zinc finger DNA-binding proteins that hierarchically regulate the expression of lineage-specific target genes, seats at four sites 136, 262, 419 and 464. The sequence also contains DOFCOREZM, EBOXBNNAPA, GT1CONSENSUS, GTGANTG10, IBOXCORE, MYBGAHV, MYBST1, POLASIG1, POLLEN1LELAT52, RAV1AAT, REALPHALGLHCB21, ROOTMOTIFTAPOX1,TAAAGSTKST1 and WBOXATNPR1 factors, the function of each regulatory element and its combined effect remain to be determined.

Analysis of ussH region reveals that there is a putative promoter sequence CATATTTTATTAAAGCGGGATGAGCTTGGTAAAATCGGTAT*CCGAGATTA (from 1620 to 1665 bp). The −10 and −35 regions are underlined, and the asterisked T represents the transcription start site. There are some functional motifs in this sequence recognized by web signal scan, including CAAT box, TATA box, BS1EGCCR, DOFCOREZM, GT1CONSENSUS, MYBCORE, ROOTMOTIFTAPOX1 and TAAAGSTKST1.

In *A. platensis* FACHB341, the SD sequences before *cpcB*, *cpcA* and *cpcH* are all GGAG. Analysis of codon usage of *A. platensis* FACHB341 shows that there

is no obvious change comparing with that of *Nostoc* sp. PCC7120 (GenBank accession no. NC_003272), *Synechocystis* sp. PCC9143 (AF068771) and *Synechococcus* sp. PCC6301 (M94218). The GC content of codon region of *A. platensis* FACHB341, *Nostoc* sp. PCC7120, *Synechocystis* sp. PCC9143 and *Synechococcus* sp. PCC6301 is 50.2, 49.3, 48.4 and 55.3%, respectively, even lower in the ussB and ussH region in *A. platensis* FACHB341.

Discussion

The 427 bp length of 5′ upstream sequence of ussB region has been obtained in cpc operon of cyanobacteria *A. platensis*, and it has been demonstrated that there are putative promoter and signal sequences in this region, and the promoter activity of this sequence had been confirmed by promoting the expression of GFP (unpublished data). The result is favorable for transformation manipulation operation in *Arthrospira* and other prokaryotes.

In cyanobacteria, the majority of -10 region of promoter sequences are highly conserved with the consensus *E. coli* sigma-70 promoter motif TANNNT (Curtis & Martin, 1994), on the contrary, it shows weak conservation with the *E. coli* sigma −70 consensus sequence (TTGACA) in −35 region, only in approximately one-half of the promoters mapped for different genera (Curtis & Martin, 1994). The functional elements of *cpcB* promoter of *Synechococcus* PCC7002 are at the position −35 (ATGACA) and −12 (TAATGT) (Bryant, 1985). The four −10 and −35 regions of the upstream sequence of *cpcB* gene are all not conserved as *E. coli* sigma-70 promoter motif. The positions of the transcription start sites are been characterized in some cyanobacteria. In *Fremyella diplosiphon*, the sequence in the range from −76 to −37 is necessary for the expression of *cpcB2A2*, and the region extending from −76 to +25 is sufficient for red-light induction of the operon. There are two protein-binding sites in the *cpcB2A2* promoter, locating in positions from −162 to −122 and −37 to +25 respectively (Casey & Grossman, 1984). In this study, the presumable transcription start sites of *cpcB* gene lie in −216 bp, −158 bp, −57 bp and −41 bp from the initiation codon, and the transcription start site of *cpcH* lies in −74 bp upstream from the initiation codon of *cpcH*. The distance between the transcription site and the initiation codon is variable in different organisms and different genes.

The result of this investigation shows that *cpcH* lies downstream of *cpcA* gene in *A. platensis* FACHB341, which codes the rod linker polypeptides, that is agreement with Jeamton (1999) and Mazel (1988). The gene encoding the rod linker polypeptides was named *cpcCCD* in *Synechocystis* PCC6803 and PCC 6714 (Kaneko, 1995, 1996; Nakajima, 2002), and *cpcCD* in *Nostoc* PCC7120 (Kaneko, 2001). The result of alignment of the sequence of *cpcH* and *cpcC* has lower similarity (result not shown), suggesting that the rod linker polypeptides are quite dispersive among the three cyanobacterial genera, *Arthrospira*, *Synechocystis* and *Nostoc*.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J. mol. Biol. 215: 403–410.

Bhat VB, Madyastha KM (2001) Scavenging of peroxynitrite by phycocyanin and phycocyanobilin from *Spirulina platensis*: Protection against oxidative damage to DNA. Biochem biophys. Res Commun. 285: 262–266.

Bryant DA, Dubbs JM, Field PI, Porter RD (1985) Expression of phycobiliprotein genes in *Escherichia coli*. FEMS Microbiol. Lett. 29: 343–349.

Casey ES, Grossman A (1984) In vivo and in vitro characterization of the light-regulated *cpcB2A2* promoter of *Fremyella diplosiphon*. J. Bact. 176: 6362–6374.

Curtis SE, Martin JA (1994) The transcription apparatus and the regulation of transcription initiation. In Bryant DA (ed.), The Molecular Biology of Cyanobacteria. Klumer Academic Press, Dordrecht, The Netherlands, pp. 613–639.

Guan Y, Guo B (2002) Inhibition activity of spirulina platensis proteins photo-immobilization biomaterial on proliferation of cancer cells. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi.19: 1–3.

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res. 27: 297–300.

Jeamton W, Chaisawadi S, Deshnium P, Castets AM, Coursin T, Tandeau de Marsac N, Tanticharoen M, Cheevadhanarak S (1996) Direct submission to GeneBank. Accession Number: Y09074.

Jeamton W, Meesapyodsuk D, Deshnium P, Castets AM, Coursin T, Tandeau de Marsac N, Tanticharoen M, Cheevadhanarak S (1999) GeneBank Accession Number: AF164139.

Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, Iriguchi M, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kohara M, Matsumoto M, Matsuno A, Muraki A (2001) Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC7120. DNA Res. 8: 205–213.

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. 3: 109–136.

Kaneko T, Tanaka A, Sato S, Kotani H, Sazuka T, Miyajima N, Sugiura M, Tabata S (1995) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. DNA Res. 2: 153–166.

Manen JF, Falquet J (2002) The *cpcB-cpcA* locus as a tool for the genetic characterization of the genus Arthrospira (Cyanobacteria): Evidence for horizontal transfer. Int. J. Syst. Evol. Microbiol. 52: 861–867.

Mazel D, Houmard J, Tandeau de Marsac N (1988) A multigene family in *Calothrix* sp. PCC7601 encodes phycocyanin, the major component of the cyanobacterial light harvesting antenna. Mol. gen. Genet. 211: 296–304.

Nakajima Y, Fujiwara S, Sawai H, Imashimizu M, Tsuzuki M (2002) A phycocyanin-deficient mutant of *Synechocystis* PCC6714 with a single-base substitution upstream of the cpc operon. Plant Cell Physiol. 42: 992–998.

Pilot TJ, Fox JL (1984) Cloning and sequencing of the genes encoding the α and β subunits of C-phycocyanin from the cyanobacterium *Agmenellum quadruplilcatum*. Proc. natl. Acad. Sci. USA 81: 6983–6987.

Pinero Estrada JE, Bermejo Bescos P, Villar del Fresno AM (2001) Antioxidant activity of different fractions of *Spirulina platensis* protein extract. Farmaco 56: 497–500.

Reddy Mc, Subhashini J, Mabipal SV (2003) C-Phycocyanin, a selective cyclooxygenase-2 inhibitor, induces apoptosis in lipopolysaccharide-stimulated RAW 264.7 macrophages. Biochem biophys. Res Commun. 304: 385–392.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual (2nd edn), Cold Spring Harbor Laboratory Press.

Yu P, Li JR, Cen PL (2002) Cloning and sequencing of the phycocyanin gene from *Spirulina maxim*a and its evolutionary analysis. J. appl. Phycol. 14: 307–310.