Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochimica et Biophysica Acta 1757 (2006) 130–134
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
http://www.elsevier.com/locate/bba

The phycocyanin-associated rod linker proteins of the phycobilisome of *Gloeobacter violaceus* PCC 7421 contain unusually located rod-capping domains

Emma Berta Gutiérrez-Cirlos [a], Bertha Pérez-Gómez [b], David W. Krogmann [c], Carlos Gómez-Lojero [b,*]

[a] *Unidad de Biomedicina FES-Iztacala UNAM Av. De los Barrios 1, Los Reyes Iztacala, Tlalnepantla, Edo de México 54090, México*
[b] *Departmento de Bioquímica, Centro de Investigación y Estudios Avanzados-IPN, Apartado Postal 14-740, 07000, Mexico D.F., Mexico*
[c] *Department of Biochemistry, Purdue University, West Lafayette, IN 47907-2063, USA*

Received 16 September 2005; received in revised form 16 December 2005; accepted 18 January 2006
Available online 10 February 2006

Abstract

Gloeobacter violaceus PCC 7421 is a unique cyanobacterium that has no thylakoids and whose genome has been sequenced [Y. Nakamura, T. Kaneko, S. Sato, M. Mimuro, H. Miyashita, T. Tsuchiya, S. Sasamoto, A. Watanabe, K. Kawashima, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, N. Nakazaki, S. Shimpo, C. Takeuchi, M. Yamada, S. Tabata, Complete Genome Structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. DNA Research 10 (2003) 137–145]. Phycobilisomes of *G. violaceus* were isolated and analyzed by SDS-PAGE followed by N-terminal sequencing. Three rod-linker subunits (CpeC, CpeD and CpeE) were identified as predicted from the genome sequence. The *cpc*C1 and *cpc*C2 genes at order locus named (OLN) glr0950 and gll 3219 encoding phycocyanin-associated linker proteins from *G. violaceus* are 56 and 55 amino acids longer at the N-terminus than the open reading frame proposed in the genome. The two amino acid extensions showed a 66% identity to one another. Also, the N-terminal extensions of these sequences were similar to domains in both the rod-capping-linker protein CpcD2 and to the C-terminus domain of the phycoerythrin-associated linker protein CpeC. These domains are not only unusual in their N-terminal location, but are unusual in that they are more closely related in sequence similarity to the C-terminus domain of the phycoerythrin-associated linker, CpeC of *G. violaceus*, than to the C-terminus domain of phycocyanin-associated linker CpcC in other cyanobacteria. These linker proteins with unique special domains are indicators of the unusual structure of the phycobilisomes of *G. violaceus.*

Keywords: *Gloeobacter violaceus*; Phycobilisome; *cpc*C gene; L_R^{35} protein; CpcD domain

1. Introduction

Phycobilisomes are supramolecular complexes of phycobiliproteins and associated linker proteins, which serve as major light-harvesting antennae in cyanobacteria and red algae. The structure of the phycobilisomes in most cyanobacteria and red algae is a hemidiscoidal shape [2–4]. *G. violaceus* is exceptional in that its phycobilisomes are not hemidiscoidal. The PBS core is amorphous and the rods are grouped as one bundle of six parallel rods of phycocyanin and phycoerythrin [5]. Genes encoding the phycobiliprotein subunits and the linker proteins of the phycobilisome of *G. violaceus* have been identified in the genome of the organism [1]. The linker proteins determine the positions of the phycobiliproteins within the phycobilisome structure [6,7]. Also, the linker proteins interact either directly or indirectly with the chromophores causing changes in their environment. These changes can modulate the spectral properties of different phycobiliprotein subassemblies [8–10]. In the amino acid sequences of the linker proteins, one sees a "conserved domain" [4,6,7] common to all of the following structures: in the PC-associated linker proteins, in the phycoerythrin-associated linkers, in the sequence repetitions in tandem (REPs) of the multidomain core-membrane linker (ApcE) [11,12] and in the linker proteins that bind the rods to the core [6,7]. The rod linker proteins (L_R^{30-35}) are linker

* Corresponding author. Tel.: +52 55 5061 3948; fax: +52 55 5061 3391.
E-mail address: cgomez@cinvestav.mx (C. Gómez-Lojero).

0005-2728/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbabio.2006.01.006

proteins associated with phycobiliproteins in the peripheral rods. Depending on the length of their phycobilisomes, cyanobacteria have one or more L_R^{30-35} proteins. In the genome of *G. violaceus*, several linkers have been identified: two open reading frames have been designated CpcC proteins, linkers associated with phycocyanin, at the OLN glr0950 (CpcC1) and gll3219 (CpcC2). Also, three genes for phycoerythrin-associated linkers, CpeC, CpeD and CpeE (at OLN glr1263, glr1264 and glr1265), and two genes encoding linker proteins that function as corks or caps in the rods of the PBS [13] were identified at OLN gsr1266 (CpcD1) and gsr1267 (CpcD2).

Homologue sequences of CpcD are found, as domains, in other proteins. For example, the N-terminal extension of ferredoxin $NADP^+$ oxidoreductase (FNR-3D) is responsible for the attachment of the protein to the distal phycocyanin in the phycobilisomes of other cyanobacteria [14,15]. Also, a CpcD-like domain is found at the C-terminus of the phycocyanin-associated rod linker [15,16]. Our group has initiated a project to characterize which genes of *G. violaceus* encoding phycobilisome proteins are expressed in order to rationalize the structure of the phycobilisomes based on its proteome. The aim of this paper is to show that CpcCs, as described in the complete genome of *G. violaceus*, are truncated. In fact, the two identified CpcC proteins have N-terminal extensions 55–56 amino acids longer which are homologous to the small linker CpcD2. We have also found that the CpcD-like domain is usually found at the C-terminus of other CpcC proteins [15].

2. Materials and methods

2.1. Cell growth and phycobilisome preparation

G. violaceus PCC 7421 was supplied by the Pasteur culture collection [17]. Cultures were grown in 50 ml of BG-11 medium, supplemented with 5 mM Mops buffer pH 7.8 [18]. The light intensity employed never exceeded 5 μmol m^{-2}. The cultures were grown for 18 days. The cells were washed once with 0.65 M Na^+/K^+ phosphate buffer pH 8 and suspended in the same buffer. The cells were disrupted twice in a French pressure cell (Thermo Spectronic) operated at 20,000 psi. The broken cells were treated with Triton X-100 (1% v/v final concentration) for 30 min and centrifuged for 10 min at 27,000×*g*. Ammonium sulfate (240 mg per ml) was added to the supernatant. The precipitate was suspended with phosphate buffer as above and was layered (2 ml per tube) on sucrose step gradients in 0.75 M Na^+/K^+ phosphate buffer pH 8. The gradients were centrifuged at 65, 000×*g* in a Beckman Ti 60 rotor for 14 h at 20 °C.

2.2. Absorption spectra

Absorption spectra were recorded with an Aminco DW2 UV-visible spectrophotometer with the OLIS conversion and OLIS software (On-Line Instrument Systems Inc. Bogart, GA). The concentrations of phycobiliproteins were estimated using the extinction coefficients for *G. violaceous* proteins of the protomer and are expressed in mM^{-1} cm^{-1}: ε_{564}^{PE} 456, ε_{564}^{PC} 140, ε_{617}^{PC} 365, ε_{650}^{PC} 64, ε_{564}^{APC} 40, ε_{617}^{APC} 145, ε_{650}^{APC} 225, [19] and using the published set of simultaneous equations [2].

2.3. N-terminal sequence analysis

L_R^{35} was prepared for N-terminal sequence analysis as follows. The SDS-PAGE was run with 150 μg of protein from the phycobilisomes in each well of a 1.5-mm thick slab gels (10% acrylamide) [20]. The proteins were then transferred to a Trans Blot (Bio Rad PVDF) membrane using the Trans Blot SD (Bio Rad) semi-Dry system. The PVDF membrane was stained with Coomassie brilliant blue R-250 and the bands were cut and sent to the Protein Separation and Analysis Laboratory at Purdue University for N-terminal sequencing.

3. Results

3.1. Phycobiliprotein spectra and stoichiometry

Fig. 1 Panel A shows the absorption spectrum of the isolated phycobilisomes of *G. violaceus*. The peak at 564 nm is due to phycoerythrin (PE), the peak at 620 is phycocyanin (PC) and the shoulder at 650 nm is allophycocyanin (AP). This type of phycobilisome contains higher ratios of both phycoerythrin and phycocyanin relative to allophycocyanin, which is expected in cells grown in dim light. The molar ratios of phycobiliproteins in the phycobilisome were 1 AP: 3 PC: 2.2 PE.

3.2. Linker proteins and their genes.

Fig. 1 Panel B shows the protein profile of purified PBS from the cyanobacterium *G. violaceus*. Three protein bands, with apparent molecular masses of 30, 33 and 35 kDa, were identified as probable linker proteins. N-terminal sequencing of the 30 and 35 kDa protein bands gave two different sequences from each band (Table 1 and Panel B of Fig. 1). Three of these sequences were unequivocally identified as CpeC, CpeD and CpeE (Table 1).

An eleven amino acids sequence, –**SVLTGDNQQRG**–, was obtained from N-terminal sequencing of the 35 kDa band. A sequence of nine amino acid residues, –**MNVLTTSSQ**–, was obtained from N-terminal sequencing of the 33-kDa band. When both of these sequences were matched with the *G.*

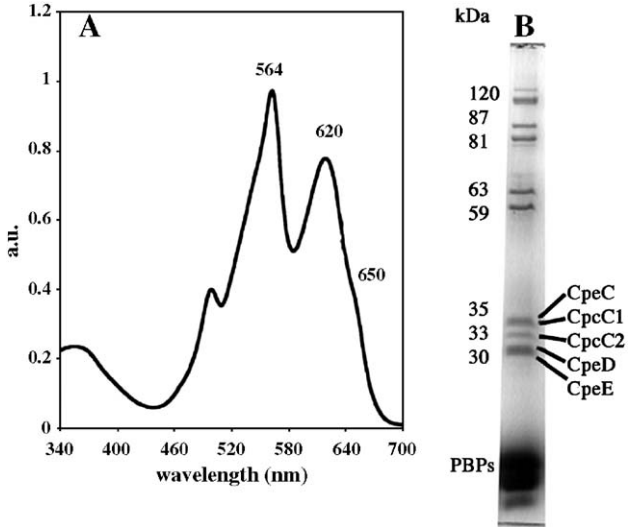

Fig. 1. Phycobilisomes of *G. violaceus* PCC 7421. Panel A. Absorption spectrum. Panel B. Protein components resolved on SDS-PAGE (10% polyacrylamide slab gel). The lane has 50 μg of protein of phycobilisome. Apparent molecular masses are indicated at the left and the rod linkers identified by N-terminal sequences are indicated at the right.

Table 1
N-terminal amino acid sequences

Cycle no.	Band of 30 kDa	pmol	Band of 33 kDa	pmol	Band of 35 kDa	pmol
1	Ser	4.9	Met	0.6	Ser/Ala	2.1/1.4
2	Leu/Thr	2.4/1.2	Asn	0.8	Ile/Val	1.4/0.8
3	Trp/Val	0.5/0.6	Val	0.7	Gly/Leu	0.8/0.6
4	Ala/Met	2.1/0.3	Leu	0.3	Thr/Pro	1.2/1.1
5	Thr	3.6	Thr	1.0	Ala/Gly	1.1/0.9
6	Glu/Gly	2.0/0.6	Thr	1.2	Ser/Asp	1.0/0.9
7	Gln/Ala	1.3/0.7	Ser	0.8	Glu/Asn	1.2/0.4
8	Ala/Val	1.6/0.5	Ser	1.0	Leu/Gln	0.9/0.5
9	Thr/Gly	2.0/0.6	Gln	0.4	Gly/Glu	1.2/0.7
10	Ile/Ala	1.3/0.8			Ile/Arg	1.0/n.q.
11	Glu/Gly	1.5/0.6			Gly/Ser	0.9/0.8

The 30 kDa band gave the N-terminal sequence –STVMTGAVGAG– and SLWATEQATIE–. The proteins identified were CpeD and CpeE. The 33-kDa band gave the N-terminal sequence –MNVLTTSSQ– used as a query in TBLSATN to find CpcC2 extension. The 35-kDa band gave the N-terminal sequence of CpeC-AIGPASELGIS–and the sequence –SVLTGDNQQRG– used as a query in TBLSATN to find CpcC1 extension.

violaceous genome, they resulted in unknown open reading frames, and the corresponding gene sequences were not found among the annotated genes.

We then performed a TBLASTN [21] analysis using the obtained sequences as queries. Both were identified as N-terminus extensions of the two existing cpcC genes of *G. violaceous*. The eleven amino acid residue peptide was ascribed to the cpcC1 gene (at glr0950; between nts: 1007494–1007662) encoding a 56 amino acid long polypeptide corresponding to the sequence: MSVLTGDNQQRGSKLFKITIALSPTLAHHPWP-GLDTH EPSQSSYSTIVSLERLLPE located at the N-terminus of the CpcC1.

The nine amino acid sequence was ascribed to gene cpcC2 (OLN gll3219, between nts: 3426726 and 3426752) encoding a 55-amino acid residue polypeptide corresponding to the sequence: MNVLTTSSQRGGKLFKVTMTLSPALSHHPWP-SLDTY EPSQNSYSVVVPLDRLLAE at the N-terminus of the CpcC2. The 55-amino acid extension was located between the intergenic region of genes CPC and cpcB2. The two new-found extensions should be added to the sequences of CpcC1 and CpcC2 predicted proteins. Those extensions showed similarity to the C-terminus domain of the phycoerythrin rod linker CpeC and to the CpcD2 protein (Fig. 2).

3.3. Similarity of the CpcC extensions

Sequence alignment of the newly identified extensions from the CpcC1 protein and the CpcC2 protein show that these extensions have a 60% identity and an 80% similarity. N-terminal extensions were aligned to the complete sequence of CpcD2 protein of *G. violaceus*. The CpcD2 protein exhibited a strong resemblance to the sequences of CpcC1 with a 52% identity and a 66% similarity and to the CpcC2 a 45% identity and a 61% similarity. With these values, we can conclude that the CpcD-like domain present at the C-terminus in the CpcCs of other cyanobacteria is in fact at the N-terminus in the CpcCs of *G. violaceus*, just as it is in the protein ferredoxin-NADPH oxidoreductase (FNR-3D) in other cyanobacteria [15,16]. The CpcD-like domain is also present, but at the C-terminus of CpeC of *G. violaceus*. Actually, this domain in the CpeC is more similar to the N-terminus of CpcC1 and CpcC2 than to CpcD2 (a 50–52% identity and a 66–70% similarity), as can be seen in Fig. 2. The CpcD-like domain is more frequently found at the C-terminus of rod linkers, as evidenced by this domain's presence at the N-terminus of the CpcCs of *G. violaceus* and FNR-3D. The meaning of this translocation remains to be identified.

Fig. 2. Amino acid sequence alignment of *G. violaceus* proteins. The comparison is between the complete sequences of CpcC2 and CpcC1, in the first two lines of sequences with the C-terminal domain, from amino acid 216 to amino acid 280 of CpeC, in the third line and with CpcD2 in the fourth line of the top cluster. The sequences shown are derived from the complete genome of *G. violaceus* [22] and aligned using ClustalX with manual adjustments. N-terminal extensions of CpcC2 and CpcC1 found in this work are shown by the black upper bar. Strictly conserved residues are marked above with an asterisk *. The dashes (-) indicate insertions/deletions and the period or colon (./:) indicate sequence similarity.

4. Discussion

4.1. Characterization of CpcCs of G. violaceus

By searching the complete genome of *G. violaceus* [22] we found that two phycocyanin-associated rod linker proteins, CpcCs, encoded as *cpc*C1 and *cpc*C2 were predicted to be 219 amino acids long. The biochemical evidence reported in this work shows very clearly that the actual lengths of the product of the genes *cpc*C1 and *cpc*C2 are 275 and 274 amino acids respectively. Two previously unrecognized N-terminal extensions of 56 and 55 amino acids were found to belong to these phycocyanin-associated rod linker proteins. The *cpc*C1 gene is now located between nts 1007497 and 1008322 encoding a protein of 275 amino acids with a calculated molecular mass of 31043 Da and a predicted isoelectric point of 6.77. In addition, the *cpc*C2, (in the negative reference DNA strand) is located between nts 3425928 and 3426752, and encodes a protein of 274 amino acids with a calculated molecular mass of 30878 and a predicted isoelectric point of 7.77.

4.2. Two domains in the CpcC proteins

The amino acid sequence comparison of the *G. violaceus* CpcC1 and CpcC2 proteins (Fig. 2), clearly shows the presence of two highly similar domains: a N-terminal domain (amino acids 1–70) and a C-terminal domain (amino acids 110–274). Between these domains (71 to 109) there are sequences of high variability, rich in polar amino acids with pI(s) of 4.11 and 3.9. These residues could form a hinge-like structure of the type present in the FNR-3D [14].

Since the N-terminal domains of the CpcC proteins exhibited a 61 to 66% similarity with CpcD2 protein, we called these domains the CpcD-like domains. The C-terminus domain of the CpcC proteins corresponds to the "conserved domain" of Bryant et al. [6].

4.3. CpcCs location and interaction with phycocyanin and phycoerythrin

The linker proteins are believed to be located mainly in the central cavity of the phycobiliprotein hexamers or trimers. The structure of a trimer of allophycocyanin with the small linker protein $L_C^{8.9}$ has been solved by X-ray diffraction studies [10]. This structure shows the linker in the cavity and in contact with one of the phycocyanobilins [10]. It is also believed that the linker proteins are interconnected, forming a skeleton-like structure within the phycobilisome. Finally, linkers vary with each disk along the length of the rod, even if the disks themselves are identical. The classical CpcCs, with the complete CpcD-like domain at the C-terminus, are found attaching the distal phycocyanin to the rod. The specificity for interaction could be inside the phycobiliproteins between the linkers. As we have shown for FNR-3D [15], the distal phycocyanin that has the L_R^{33} linker inside has a higher affinity for the CpcD-like domain of FNR-3D, while the phycocyanin with L_{RC}^{27} has a lower affinity for FNR-3D.

It is clear that special linkers are needed at the interface of the phycobiliproteins. The interface between phycocyanin and phycoerythrin is relevant to this discussion. The special CpcCs of *G. violaceus* might be at the interface between these two phycobiliproteins. The conservative domain might interact with phycocyanin and the CpcD-like domain may interact with phycoerythrin. The presence of a CpcD-like domain at the N-terminus of the CpcC proteins that is more similar to the phycoerythrin-associated linker protein and a C-terminus domain that is more similar to the phycocyanin-associated linker protein are consistent with that localization.

Further experiments on the partial dissociation of the *G. violaceus* phycobilisomes are needed to elucidate the localization of the CpcC proteins in these complex light harvesting antennae.

Finally, it is important to mention the special sequence HHPWP present in the CpcD-like domain of CpcCs (gray bar in Fig. 2). This sequence is not found in either CpeC or CpcD2. The presence of the tryptophan between those rigid prolines could be important in the tuning of the chromophore for optimal light harvesting.

Acknowledgements

We are greatly indebted to Diego González-Halphen for his editorial assistance and stimulating discussions. We also thank Mary Bower of Purdue University for her expert work in N-terminal sequencing, Jorge Zarco Mendoza for his technical assistance and Cristina Díaz-T for her secretarial assistance. We also appreciate critical reading of the manuscript by Loretta Krogmann and Michelle Krogmann.

This study was supported by Grant SEP-CONACYT-México (ref. No. 46416).

References

[1] Y. Nakamura, T. Kaneko, S. Sato, M. Mimuro, H. Miyashita, T. Tsuchiya, S. Sasamoto, A. Watanabe, K. Kawashima, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, N. Nakazaki, S. Shimpo, C. Takeuchi, M. Yamada, S. Tabata, Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylacoids, DNA Res. 10 (2003) 137–145.

[2] D.A. Bryant, G. Guglielmi, Tandeau de Marsac, A.M. Castets, G. Cohen Bazire, The structure of cyanobacterial phycobilisomes: a model, Arch. Microbiol. 123 (1979) 113–127.

[3] A.N. Glazer, Light guides, J. Biol. Chem. 264 (1989) 1–4.

[4] W.A. Sidler, Phycobilisome and phycobiliprotein structures, in: D.A. Bryant (Ed.), The Molecular Biology of Cyanobacteria, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994, pp. 139–216.

[5] G. Guglielmi, G. Cohen-Bazire, D.A. Bryant, The structure of *Gloeobacter violaceus* and its phycobilisomes, Arch. Microbiol. 129 (1981) 181–189.

[6] D.A. Bryant, V.L. Stirewalt, M. Glauser, G. Frank, W. Sidler, H. Zuber, A small multigene family encodes the rod-core linker polypeptides of *Anabaena* sp PCC7120 phycobilisomes, Gene 107 (1991) 91–99.

[7] M. Glauser, W. Sidler, G. Frank, H. Zuber, Studies on the rod substructure of the phycobilisome from the cyanobacterium *Mastigocladus laminosus*, in: M. Baltscheffsky (Ed.), Current Research in Photosynthesis, vol II, 1990, pp. 89–92.

[8] L. Gottschalk, R Fischer, F. Lottspeich, H. Scheer, Origin of the red shifted absorption in phycocyanin 632 from *Mastigocladus laminosus*, Photochem. Photobiol. 54 (1991) 283–288.

[9] L. Gottschalk, F. Lottspeich, H. Scheer, Reconstitution of allophycocyanin from *Mastigocladus laminosus* with isolated linker polypeptide, Photochem. Photobiol. 58 (1993) 761–767.
[10] W. Reuter, G. Wiegand, R. Huber, M.E. Than, Structural analysis at 2.2 Å of orthorhombic crystals presents the asymmetry of the allophycocyanin-linker complex, AP.$L_C^{7.8}$ from phycobilisomes of *Mastigocladus laminosus*, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 1363–1368.
[11] D.A. Bryant, Genetic analysis of the phycobilisome biosynthesis, assembly, structure and function in the cyanobacterium *Synechococcus* sp PCC 7002, in: S.E. Stevens Jr., D.A. Bryant (Eds.), Light-Energy Transduction in Photosynthesis: Higher Plants and Bacterial Models, American Society of Plant Physiologists, Rockville, 1988, pp. 62–90.
[12] J. Houmard, V. Capuano, M.V. Colombano, T. Coursin, N. Tandeau de Marsac, Molecular characterization of the terminal energy acceptor of cyanobacterial phycobilisomes, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 2152–2156.
[13] R. de Lorimier, D.A. Bryant, S.E. Stevens Jr., Genetic analysis of a 9 kDa phycocyanin/associated linker polypeptide, Biochim. Biophys Acta 1019 (1990) 29–41.
[14] W.M. Schluchter, D.A. Bryant, Molecular characterization of ferredoxin-NADP+ oxidoreductase in cyanobacteria: cloning and sequence of the *pet*H gene of *Synechococcus* sp PCC 7002 and studies on the gene product, Biochemistry 31 (1992) 3092–3102.
[15] C. Gómez-Lojero, B. Pérez-Gómez, G. Shen, W.M. Schluchter, D.A. Bryant, Interaction of ferredoxin:$NADP^+$ oxidoreductase with phycobilisomes and phycobilisome substructures of the cyanobacterium *Synechococcus* sp strain PCC 7002, Biochemistry 42 (2003) 13800–13811.
[16] R. de Lorimier, G. Guglielmi, D.A. Bryant, S.E. Stevens Jr., Structure and mutation of a gene encoding 33000-molecular weight phycocyanin-associated linker polypeptide, Arch. Microbiol. 153 (1990) 541–549.
[17] R. Rippka, J. Waterbury, G. Cohen-Bazire, A cyanobacterium which lacks thylakoids, Arch. Microbiol. 100 (1974) 419–436.
[18] R. Rippka, J. Deruelles, J.B. Waterbury, M. Herdman, R.Y Stanier, Generic assignments, strain histories and properties of pure cultures of cyanobacteria, J. Gen. Microbiol. 111 (1979) 1–61.
[19] D.A. Bryant, G. Cohen-Bazire, A.N. Glazer, Characterization of the biliproteins of *Gloeobacter violaceus* chromophore content of a cyanobacterial phycoerythrin carrying phycourobilin chromophore, Arch. Microbiol. 129 (1981) 190–198.
[20] H. Schägger, G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, Anal. Biochem. 166 (1987) 368–379.
[21] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[22] Y. Nakamura, T. Kaneko, S. Sato, M. Mimuro, H. Miyashita, T. Tsuchiya, S. Sasamoto, A. Watanabe, K. Kawashima, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, N. Nakazaki, S. Shimpo, C. Takeuchi, M. Yamada, S. Tabata, Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids (Supplement), DNA Res. 10 (2003) 181–201.