*Proc. Natl. Acad. Sci. USA*
Vol. 87, pp. 2152–2156, March 1990
Biochemistry

# Molecular characterization of the terminal energy acceptor of cyanobacterial phycobilisomes

(*Calothrix* sp. PCC 7601/*apcE* sequence/$L_{CM}$ polypeptide/phycobiliprotein/secondary structure)

JEAN HOUMARD*, VÉRONIQUE CAPUANO, MARCUS V. COLOMBANO, THÉRÈSE COURSIN, AND NICOLE TANDEAU DE MARSAC

Physiologie Microbienne, Centre National de la Recherche Scientifique Unité de Recherche Associée 1129, Département de Biochimie et Génétique Moléculaire, Institut Pasteur, 28 rue du Dr Roux, F-75724 Paris Cedex 15, France

*Communicated by Lawrence Bogorad, December 28, 1989*

ABSTRACT **Cyanobacteria harvest light energy through multimolecular structures, the phycobilisomes, regularly arrayed at the surface of the photosynthetic membranes. Phycobilisomes consist of a central core from which rods radiate. A large polypeptide ($L_{CM}$, 75–120 kDa) is postulated to act both as terminal energy acceptor and as a linker polypeptide that stabilizes the phycobilisome architecture. We report here the characterization of the gene (*apcE*) that encodes this $L_{CM}$ polypeptide in *Calothrix* sp. PCC 7601. It is located upstream from the genes encoding the major components of the phycobilisome core (allophycocyanin) and is part of the same operon. The deduced amino acid sequence shows that the N-terminal region of $L_{CM}$ shares homology with the other phycobiliprotein subunits and thus constitutes the chromoprotein domain. The other part of the molecule is made up of four repeated domains that are highly homologous to the N-terminal regions of the phycocyanin rod linker polypeptides. The predicted secondary structure of the different domains of the $L_{CM}$ is discussed in relation to the different roles and properties of this large molecule.**

Cyanobacteria, like all photosynthetic organisms, adjust their pigment content to the available light intensity, synthesizing more pigment under low than under high light intensities. In addition, some of the cyanobacteria that synthesize phycoerythrin, a phycobiliprotein that absorbs green light, can modify their phycobiliprotein content in response to changes in the light wavelengths received during cell growth. This phenomenon is known as complementary chromatic adaptation (for a review, see ref. 1). These adaptation processes mainly affect phycobilisomes, which constitute the light-harvesting antennae of cyanobacteria. These fanlike structures are made up of a central core from which six rods radiate (2). The cores are mainly composed of allophycocyanin ($\alpha^{AP}\beta^{AP}$) with which minor components, $\alpha^{APB}$, $\beta^{18.3}$, $L_C^{7.8}$, and $L_{CM}$, are associated (3–7). Among them, $L_{CM}$ plays a key role since it acts as a terminal energy acceptor of the phycobilisome, as a linker polypeptide that contributes to the phycobilisome architecture, and as a molecule that connects ("anchors") the phycobilisome to the thylakoid membrane. There are two copies of $L_{CM}$ per phycobilisome (for a review, see ref. 8).

As part of ongoing studies on the photoregulation of gene expression, we have isolated most of the genes that encode the phycobilisome components of *Calothrix* sp. PCC 7601, a cyanobacterium that belongs to the family of the chromatic adapters (for reviews, see refs. 9 and 10). In particular, the *apcA1* ($\alpha^{AP}$), *apcB1* ($\beta^{AP}$), *apcC* ($L_C^{7.8}$), and *apcD* ($\alpha^{APB}$) genes, which encode core components, have been characterized (11, 12). The first three genes are adjacent on the chromosome and form the *apcA1B1C* operon. During the characterization of these genes, we observed that a gene, *apcE*, located upstream from *apcA1* was also part of this operon and probably corresponded to $L_{CM}$ (12). In this report, we describe the complete characterization of the *apcE* gene, the product of which is essential for the building up of functional phycobilisomes.†

## MATERIALS AND METHODS

**Culture Conditions and Isolation of DNA and RNA.** Culture conditions for *Calothrix* sp. strain PCC 7601 (= *Fremyella diplosiphon* UTEX 481, hereafter designated *Calothrix* 7601), as well as chromatic illuminations and total DNA and total RNA purifications, have been described (13). Small-scale plasmid extractions were performed on overnight cultures of *Escherichia coli* as described (14). Standard procedures were used for large-scale plasmid preparation, DNA fragment isolation, and agarose gel electrophoresis (15).

**Genomic Library Construction and Hybridization.** The partial DNA library was constructed by ligation of *Hin*dIII DNA fragments [9–12 kilobases (kb)] into the *Hin*dIII site of plasmid pTZ18R. Standard methods were used for *in situ* colony hybridization (15). Southern and Northern transfers and hybridization were performed as described (16).

**DNA Sequence Analysis.** Large DNA fragments were subcloned into either pTZ18R or pTZ19R (Pharmacia). Overlapping clones were obtained by using the Cyclone system from IBI (distributed by Genofit) adapted for single-stranded DNAs of the pTZ18R subclones (17). DNA sequence analysis was performed by the chain-termination method (18) on single-stranded DNA templates. S1 nuclease mapping and primer extension experiments were performed as described (19). Computer analysis of the DNA sequence information obtained was carried out using the programs developed by the Unité d'Informatique Scientifique of the Institut Pasteur (Paris).

## RESULTS AND DISCUSSION

**Cloning of the *apcE* Gene.** We previously reported (12) that *apcA1*, *apcB1*, *apcC*, and the 3′ end of *apcE* were located on a 3.7-kb *Eco*RI DNA fragment. Southern blots of total *Calothrix* 7601 genomic DNA probed with the *Eco*RI–*Eco*RV fragment carrying part of the 3′ end of *apcE* (see Fig. 1*A*) revealed that a 10-kb *Hin*dIII fragment must contain the entire *apcEA1B1C* operon (data not shown). Using the same probe, we screened a partial *Hin*dIII gene library (see *Ma-*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: nt, nucleotide(s).

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M31224).

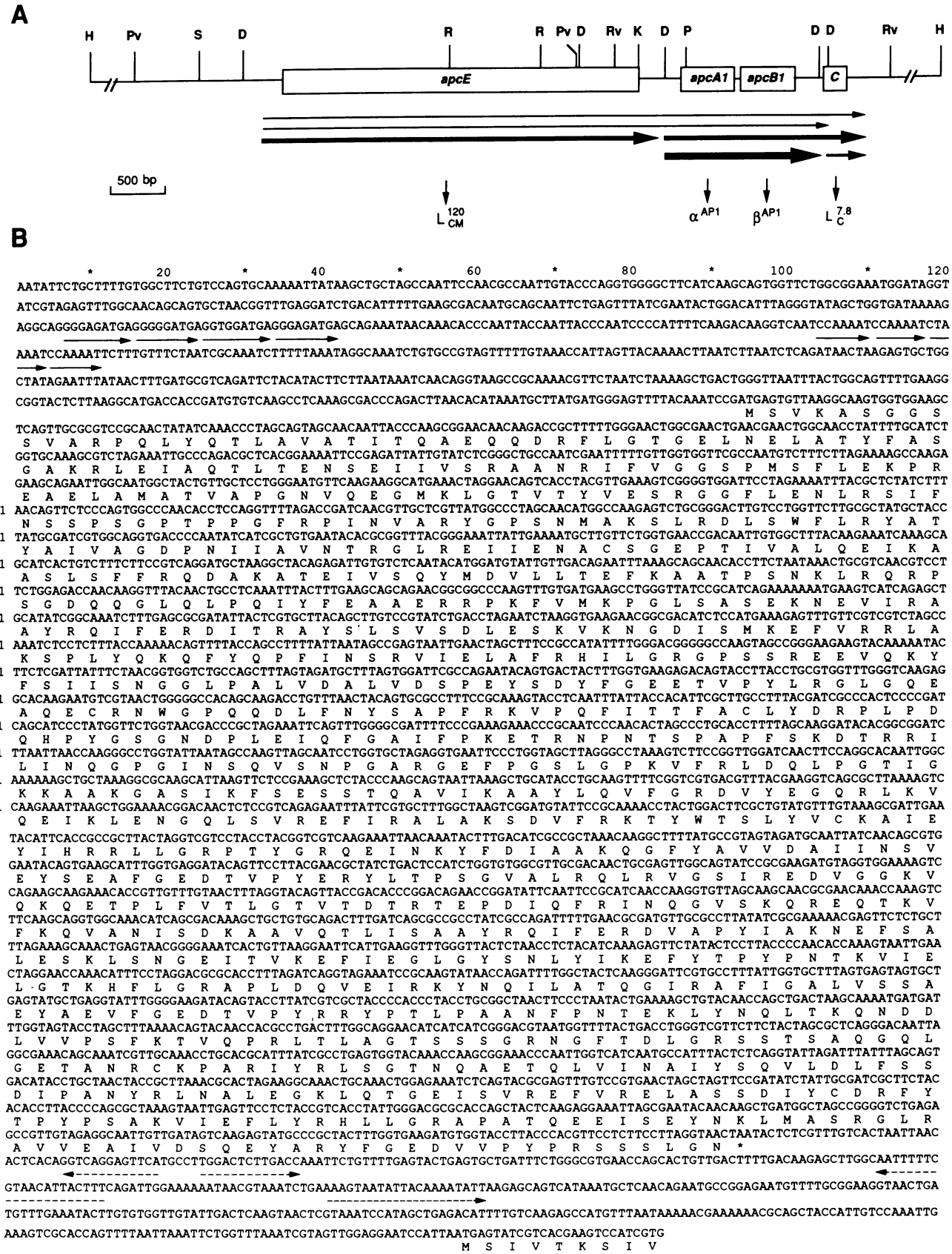


FIG. 1. (*A*) Physical map of the *apcEA1B1C* operon of *Calothrix* 7601. Boxes represent the protein coding sequences and contain gene designations. Horizontal arrows indicate the identified transcripts originating from this DNA fragment, their thickness representing their relative abundance. The encoded gene products are given below. Restriction sites: D, *Dra* I; H, *Hin*dIII; K, *Kpn* I; P, *Pst* I; Pv, *Pvu* II; R, *Eco*RI; Rv, *Eco*RV; S, *Ssp* I. bp, Base pairs. (*B*) Nucleotide sequence and deduced amino acid sequence of the *apcE* gene. Numbering of nucleotides starts at the *Ssp* I site. Standard single-letter amino acid code is used. This sequence, plus additional flanking sequences, has been deposited in the EMBL/GenBank data base. Arrows indicate direct repeats and broken arrows indicate palindromic sequences.

*terials and Methods*) and isolated a recombinant plasmid, pPM75. Further restriction analysis of this DNA fragment allowed us to conclude that most of the *apcE* gene was located on a 4.4-kb *Pvu* II fragment. This fragment was subcloned into the *Sma* I site of pTZ18R, giving rise to pPM86 and pPM87, which differed only by the orientation of the insert. A physical map of the region is represented in Fig. 1*A*.

**Nucleotide Sequence of the *apcE* Gene.** The complete nucleotide sequence of the noncoding (mRNA-like) strand of *apcE* is shown in Fig. 1*B*. It was determined on both strands by sequencing overlapping deletions of single-stranded DNAs generated with the Cyclone system, using a sequencing strategy based on the method of Dale *et al.* (20). An open reading frame of 1080 codons was found, which is preceded by a putative ribosome binding site (GGAG) located 14 base pairs (bp) upstream from the initiation codon AUG (see ref. 21 for a discussion on cyanobacterial ribosome binding sites). The *apcE* and *apcA1* genes are separated by 442 bp. In this intergenic region, two hypothetical stem-and-loop structures can be drawn. The first one starts 32 nucleotides (nt) downstream from the *apcE* stop codon, with a stem 13 nt long and a loop of 5 nt. Its thermodynamic stability ($\Delta G$) is -25.6 kcal/mol, calculated according to the method of Salser (22) as modified by Cech *et al.* (23). The second stem-and-loop starts 74 nt downstream from the first secondary structure, has a stem 21 nt long, a loop of 28 nt, and a $\Delta G$ of -16.2 kcal/mol. This structure is 43 nt upstream from the +1 position of the *apcA1B1C* transcript that we previously mapped (12).

Upstream from *apcE* lie two regions of repeated sequences. Four adjacent direct repeats of a consensus motif GGKRGATGA (K = G or T; R = A or G) and four of a consensus CYAAAAT (Y = C or T) begin 447 and 350 nt, respectively, upstream from the ATG of the *apcE* gene. Two groups of tandem repeats, each composed of eight adjacent 7-nt-long direct repeats, were already observed upstream from the *thrB* gene of *Calothrix* 7601 (24). In fact, the occurrence of such clusters of adjacent direct repeats is quite frequent within the *Calothrix* 7601 genome. From hybridization experiments, it has been deduced that probably about 100 copies of three other short tandemly repeated sequences (CCCCART, TTKGTCA, and CAACAGT) are present per genome (D. Mazel, J.H., and N.T.M., unpublished data). A similar observation has been reported for *Nostoc* sp. PCC 7906 (25). Two families of tandem direct repeats (one of 9 nt and one of 7 nt) were detected in the *petC–petA* intergenic region and they occurred in multiple copies within the *Nostoc* genome (25). However, in contrast to the data obtained with *Nostoc*, in no instance was transcription of the *Calothrix* tandem repeats detected (D. Mazel, personal communication). At present, no function can be assigned to these direct repeats, though it is tempting to speculate that they could play a role in the regulation of gene expression, being, for example, operator sites that directly regulate transcription of the downstream genes, or targets involved in DNA compaction through interactions with HU-like proteins.

**Transcriptional Analysis and Gene Organization.** Northern analyses were performed using total RNA extracted from both red- and green-light-grown cells. Similar results were obtained whatever the light wavelengths used for cell growth. When the 300-bp-long *Eco*RI–*Pvu* II DNA fragment located at the 3′ end of the *apcE* gene was used as probe, a 3.8-kb major mRNA was detected (data not shown). The larger transcripts, previously identified with an *apcA1B1* DNA probe (see figure 5 of ref. 12), were also revealed by the *apcE* probe but appeared to be about 10 times less abundant. Six mRNA species thus arise from the *apcEA1B1C* operon (Fig. 1*A*). Whether these segmented transcripts result from transcription initiation at different promoters or from processing of the larger species remains to be established. However, since we were able to detect a 5′ end in front of *apcE* (data not shown) as well as in front of *apcA1* (12), we favor the hypothesis of the existence of two promoters, one in front of the operon and one located within the *apcE–apcA1* intercistronic region.

To our knowledge, *apcE* genes have been studied in five organisms, *Nostoc* sp. strain MAC (26), *Cyanophora paradoxa* (27), *Synechococcus* sp. strains PCC 7002 (27) and PCC 6301 (10), and *Calothrix* 7601 (this work). From these studies it appears that, although the sequences are very homologous, the physical organization of the *apcA*, *apcB*, *apcC*, and *apcE* genes differs. Except in *Synechococcus* sp. PCC 7002, the *apcE* genes map just upstream from the *apcA* genes. On the other hand, in contrast to the cyanobacterial genomic arrangement, the *apcC* gene of *C. paradoxa* is nuclear-encoded, whereas the *apcEAB* cluster is part of the cyanellar genome. Recent results indicate that the *apcE* gene of *Synechococcus* sp. PCC 6301, although adjacent, is not cotranscribed with the *apcABC* operon (ref. 10; V.C. and J.H., unpublished data). A differential stability of segments of large transcripts thus appears not to be a ubiquitous mechanism used by cyanobacteria to produce the phycobilisome components in the appropriate molar ratio.

**The *apcE* Gene Product.** The primary translation product of *apcE* is predicted to be a 120-kDa polypeptide of 1080 residues (Fig. 1*B*). By analogy with the N-terminal sequences of the $L_{CM}$ of a cyanobacterial *Nostoc* sp. strain and of the eukaryotic alga *Porphyridium cruentum*, determined by Gantt *et al.* (28), that of the *Calothrix* 7601 $L_{CM}$ is SVKASGGSSVARPQLYQTLAV. This 21-residue sequence is 95% identical with the corresponding *Nostoc* sequence. As already observed for the other core components of *Calothrix* 7601 phycobilisomes (the *apcA1*, *apcB1*, *apcA2*, *apcD*, and *apcC* gene products), the N-terminal methionine residue of the *apcE* gene product is likely to be posttranslationally removed (11, 12). In contrast, rod components, both phycobiliproteins and linker polypeptides, begin with an unprocessed methionine except for the gene products of the *cpc3* operon, the expression of which only occurs under sulfur-limiting conditions (29).

The amino acid sequence of the $L_{CM}$, as deduced from the nucleotide sequence, presents some peculiar features. Its N-terminal region, residues 18–237, shares homology with phycobiliprotein subunits. Two domains of homology, separated by an insertion 68 residues long, clearly emerged from the alignment shown in Fig. 2*A*. However, the $L_{CM}$ sequence is not only 49% and 53% homologous to $\alpha^{AP1}$ and $\beta^{AP1}$, respectively, but also shares 45–48% homology with the $\alpha^{AP2}$, $\alpha^{APB}$, $\alpha^{PE}$, and $\beta^{PE}$ subunits, as well as with the three types of $\alpha^{PC}$ and $\beta^{PC}$ of *Calothrix* 7601. Although a very high conservation exists, especially with $\beta$-type subunits (up to 80% with $\beta^{PC1}$ and $\beta^{PC2}$), around the chromophore attachment site common to all phycobiliproteins (residues 82–104), Cys-85 of the phycobiliprotein subunits, which is involved in the thioether linkage of the chromophore, is replaced by a serine residue at 157 in the $L_{CM}$ sequence. The closest cysteine is at 195 in the $L_{CM}$ sequence (Cys-194 of the $L_{CM}$ after removal of the N-terminal methionine) and is likely to be the chromophore attachment site, by analogy with the amino acid sequence of the chromopeptide isolated from the $L_{CM}$ of *Synechococcus* sp. PCC 6301 (D. J. Lundell and A. N. Glazer, cited in ref. 27). Although Cys-194 is located in a region of rather poor homology (Fig. 2*A*), the highly conserved regions that precede it, which are common to all the phycobiliprotein subunits, most probably have a related tertiary structure. This structure was shown, from x-ray analysis of phycocyanin subunits, to be a long $\alpha$-helix (helix E; ref. 31). Furthermore, residue 125 of the phycocyanin subunits, which corresponds to Cys-194 of the $L_{CM}$, is part of helix F and is close to, and most probably interacts with, the

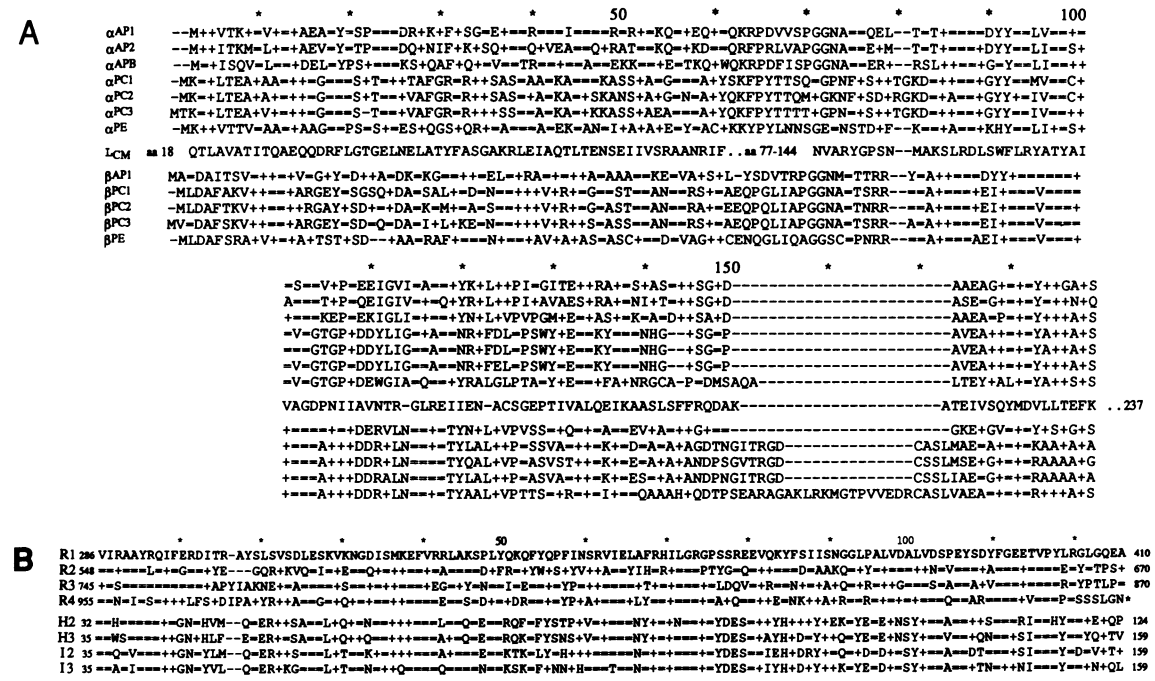


FIG. 2. (*A*) Alignment of the N-terminal domain of $L_{CM}$ with the known sequences of phycobiliprotein subunits (AP, allophycocyanin; PC, phycocyanin; PE, phycoerythrin) from *Calothrix* 7601. Numbering above the alignment refers to that of the phycobiliprotein subunits (9). =, Identity; +, conservative substitution (A, G; D, E; I, L, V, M, F; S, T, C, N, Q; K, R); −, gap introduced to maximize homology; aa, amino acids. (*B*) Alignment of the four repeated elements (R1–R4) of $L_{CM}$ with the $L_R^{PC}$ sequences of *Calothrix* 7601. Numbering above the alignment is arbitrary. Numbers at the ends of each sequence refer to its precise location within the sequence. The H2 (*cpcH2*) and I2 (*cpcI2*) sequences are from ref. 30; the H3 (*cpcH3*) and I3 (*cpcI3*) sequences are from refs. 2 and 9.

phycocyanobilin chromophore. It is thus possible that, instead of being linked by ring A, the linear tetrapyrrole would be linked through ring D to Cys-194 of the $L_{CM}$ but would maintain similar interactions with the polypeptide backbone. Alternatively, the chromophore could be linked by ring A, but then its spatial orientation would be slightly different. From recent data (32, 33), it seems that all singly linked bilins of phycobiliproteins are attached through the A ring, but D-ring linkages have also been described for doubly linked bilins (34). The shift towards longer wavelengths observed for $L_{CM}$ might be the consequence of the difference between $L_{CM}$ and the phycobiliprotein subunits with regard to the chromophore linkage and/or to its spatial orientation (35). Structural studies of the chromophore-containing peptide of $L_{CM}$ would be needed to elucidate that point.

Following the N-terminal domain, four repeated sequences were detected that are highly homologous with the sequences of the large rod linker polypeptides associated with the *Calothrix* 7601 phycocyanins ($L_R^{PC}$). Fig. 2*B* shows these four repeats aligned with one another and with the N-terminal domains of the $L_R^{PC}$ polypeptides (*cpcH* and *cpcI* gene products). Sequence homology between the four repeats ranges from 61% to 68% and it varies from 55% to 61% between repeat 1 and the four $L_R^{PC}$ sequences, for example. Closer examination of the repeat sequences shows that the first half of each repeat (VIR . . . INSR, for repeat 1) is homologous to the second half (VIE . . . VPYL, for repeat 1). The homology is maximal (43%) between the two halves of repeat 1 and minimal (32%) between those of repeat 2. Similarly, homologies also exist between the two halves of the N-terminal regions of the $L_R^{PC}$ polypeptides. Secondary structure homologies are even more pronounced and were detected by using the hydrophobic cluster analysis (HCA) method (36). Fig. 3 shows, as an example, the comparison of the HCA plots of the two halves of repeat 1. From this analysis, it appears that each half of the repeats would fold into four successive $\alpha$-helices.

In most cyanobacteria, from existing models of core substructures, two cylinders made up of four trimers lie antiparallel to each other and constitute the basis of the core, while a third cylinder is located on top of them. In these models, the $L_{CM}$ is part of the central $(\alpha^{AP}\beta^{AP})_2\beta^{18.3}L_{CM}$ trimer and two such trimers exist per phycobilisome, one in each of the two basal cylinders (2–5, 8, 27). Unfortunately, x-ray data so far

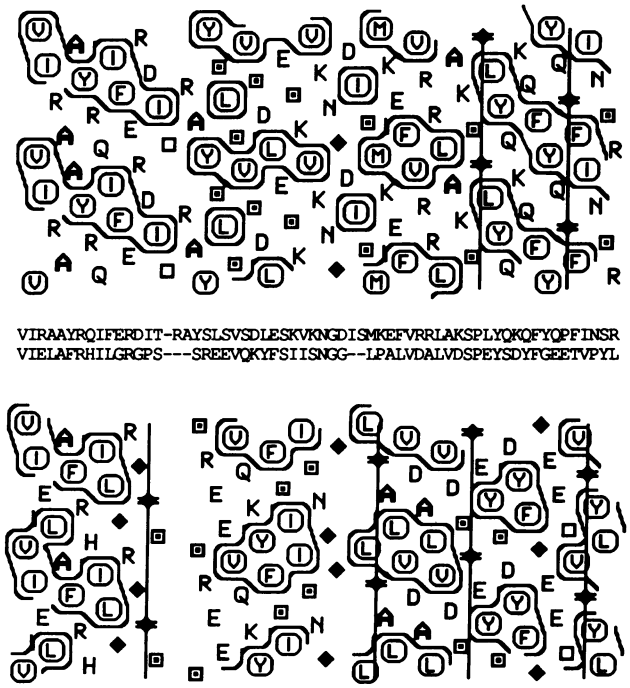


FIG. 3. HCA plots (36) and deduced alignment of the two halves of repeat 1 (amino acids 286–346 and 347–403) from the *Calothrix* 7601 $L_{CM}$. Dashes indicate gaps introduced in the sequence.

obtained on phycobilisome components originate from phycocyanin oligomers crystallized in the absence of linker polypeptides. It is therefore difficult to speculate on putative interactions between the internal repeats of the $L_{CM}$ and the different trimers and cylinders that constitute the core. In addition, although postulated to have phycobilisome cores also made up of three cylinders, cyanelles of *C. paradoxa* have an $L_{CM}$ with only three repeats (27). The $L_{CM}$ present in the three-cylinder cores of *Synechococcus* sp. PCC 7002 would also have only three repeats (D. A. Bryant, personal communication). Differences might thus exist between phycobilisome core substructures among strains that, at present, are all considered as having the same classical three-cylinder core.

Because of the very high homologies shared by the repeats, it is likely that each repeat interacts with similar structures. However these structures are spatially located at various distances from the center of the core and consequently the size of the region between the different repeat domains must vary. Indeed, the spacing arms are 45–143 amino acids long and, as expected, share much less sequence homology than the repeats. Further, the amino acid compositions of the repeats and of the spacing arms differ. Indeed, the loops contain two times more polar than hydrophobic amino acids, while the ratio of polar versus hydrophobic amino acids is about 1.2 within the repeats. Moreover, the content of proline and glycine residues, known to be often present in loops, is much higher in the spacing arms than in the repeats. Thus, the repeats seem to be more regularly folded than the spacing arms, they probably interact tightly with the phycobiliprotein trimers, and very few of their residues must be exposed to the solvent.

Besides its role as a terminal energy acceptor and as a linker involved in the stability of the phycobilisome, $L_{CM}$ is often referred to as the anchor polypeptide that links the phycobilisome with the photosystem II components, which are embedded in the photosynthetic membrane. From sequence analysis, we were unable to detect any characteristic transmembrane elements. Further, no homology was detected at the C terminus with the known consensus sequence for acylation sites (37). However, repeat 4, which constitutes the end of the $L_{CM}$, could represent a domain whereby the $L_{CM}$ interacts with components of the thylakoid membrane. Alternatively, this additional domain could strengthen the interactions between the two basal cylinders and the upper one. It might be of interest to examine the *apcE* genes of different strains and to see whether the number of repeats could be related to core substructure stability, to tighter interactions of the phycobilisomes with the thylakoid membrane, and/or to the presence of an *apcA2* gene ($\alpha^{AP2}$) in the strain. On another hand, the N-terminal domain carries the chromophore, the terminal energy acceptor, and thus has to be close to the thylakoid membrane. When compared to the phycobiliprotein subunits, this domain has two additional sequences: the N-terminal extension (residues 1–17) and the insertion (residues 77–144) that just precedes the two helices with which the chromophore is in contact. Thus, interaction between phycobilisomes and photosystems might also well be mediated through one or both of these sequences.

It has been reported that the $L_{CM}$ of *Anacystis nidulans* (*Synechococcus* sp. PCC 7942) is glycosylated (38). The consensus sequence Asn-Xaa-Ser/Thr is the most commonly recognized site for glycosylation (39). While glycosylation of the *Calothrix* 7601 $L_{CM}$ has not yet been studied, such motifs occur four times in its sequence (residues 130–132, 424–426, 476–478, and 735–737). All these sites are located in presumptive loops, as expected for glycosylation sites. However, covalent attachment of carbohydrates to the $L_{CM}$ might be questioned since the putative glycosylation sites are not conserved in the $L_{CM}$ sequences known so far (27) and since only two such presumptive sites exist in the $L_{CM}$ sequence of *Synechococcus* sp. PCC 6301 (V.C. and J.H., unpublished data).

We thank Dr. D. A. Bryant for sharing data with us prior to publication, Dr. J. P. Mornon for helpful discussions, and Dr. M. Herdman for careful reading of the manuscript. This work was supported by the Centre National de la Recherche Scientifique (URA 1129) and by the Institut Pasteur.

1. Tandeau de Marsac, N. (1983) *Bull. Inst. Pasteur* **81**, 201–254.
2. Bryant, D. A., Guglielmi, G., Tandeau de Marsac, N., Castets, A.-M. & Cohen-Bazire, G. (1979) *Arch. Microbiol.* **123**, 113–127.
3. Anderson, L. K. & Eiserling, F. A. (1986) *J. Mol. Biol.* **191**, 441–451.
4. Gantt, E. (1981) *Annu. Rev. Plant Physiol.* **32**, 327–347.
5. Gingrich, J. C., Lundell, D. J. & Glazer, A. N. (1983) *J. Cell. Biochem.* **22**, 1–14.
6. Glazer, A. N., Lundell, D. J., Yamanaka, G. & Williams, R. C. (1983) *Ann. Microbiol. (Paris)* **134B**, 159–180.
7. Glazer, A. N. (1985) *Annu. Rev. Biophys. Biophys. Chem.* **14**, 47–77.
8. Glazer, A. N. (1987) in *The Cyanobacteria*, eds. Fay, P. & Van Baalen, C. (Elsevier, Amsterdam), pp. 69–94.
9. Tandeau de Marsac, N., Mazel, D., Damerval, T., Guglielmi, G., Capuano, V. & Houmard, J. (1988) *Photosynth. Res.* **18**, 99–132.
10. Tandeau de Marsac, N., Mazel, D., Capuano, V., Damerval, T. & Houmard, J. (1990) in *Molecular Biology of Membrane-Bound Complexes in Phototrophic Bacteria*, ed. Drews, G. (Plenum, New York), in press.
11. Houmard, J., Capuano, V., Coursin, T. & Tandeau de Marsac, N. (1988) *Mol. Microbiol.* **2**, 101–107.
12. Houmard, J., Capuano, V., Coursin, T. & Tandeau de Marsac, N. (1988) *J. Bacteriol.* **170**, 5512–5521.
13. Mazel, D., Guglielmi, G., Houmard, J., Sidler, W., Bryant, D. A. & Tandeau de Marsac, N. (1986) *Nucleic Acids Res.* **14**, 8279–8290.
14. Holmes, D. S. & Quigley, M. (1981) *Anal. Biochem.* **114**, 193–197.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Damerval, T., Houmard, J., Guglielmi, G., Csiszàr, K. & Tandeau de Marsac, N. (1987) *Gene* **54**, 83–92.
17. Mazel, D., Houmard, J. & Tandeau de Marsac, N. (1988) *Mol. Gen. Genet.* **211**, 296–304.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
19. Csiszàr, K., Houmard, J., Damerval, T. & Tandeau de Marsac, N. (1987) *Gene* **60**, 29–37.
20. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31–40.
21. Tandeau de Marsac, N. & Houmard, J. (1987) in *The Cyanobacteria*, eds. Fay, P. & Van Baalen, C. (Elsevier, Amsterdam), pp. 251–302.
22. Salser, W. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **77**, 985–1010.
23. Cech, T. R., Tanner, N. K., Tinoco, I., Jr., Weir, B. R., Zuker, M. & Perlman, P. S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3903–3907.
24. Parsot, C. & Mazel, D. (1987) *Mol. Microbiol.* **1**, 45–52.
25. Kallas, T., Spiller, S. & Malkin, R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5794–5798.
26. Zilinskas, B. A. & Howell, D. A. (1987) in *Progress in Photosynthesis Research*, ed. Biggins, J. (Nijhoff, Dordrecht, F.R.G.), pp. 161–164.
27. Bryant, D. A. (1988) in *Light-Energy Transduction in Photosynthesis: Higher Plant and Bacterial Models*, eds. Stevens, S. E., Jr., & Bryant, D. A. (Am. Soc. Plant Physiol., Rockville, MD), pp. 62–90.
28. Gantt, E., Cunningham, F. X., Jr., Lipschultz, C. A. & Mimuro, M. (1988) *Plant Physiol.* **86**, 996–998.
29. Mazel, D. & Marlière, P. (1989) *Nature (London)* **341**, 245–248.
30. Lomax, T. L., Conley, P. B., Schilling, J. & Grossman, A. R. (1987) *J. Bacteriol.* **169**, 2675–2684.
31. Schirmer, T., Huber, R., Schneider, M., Bode, W., Miller, M. & Hackert, M. L. (1986) *J. Mol. Biol.* **188**, 651–676.
32. Lagarias, J. C., Klotz, A. V., Dallas, J. L., Glazer, A. N., Bishop, J. E., O'Connell, J. F. & Rapoport, H. (1988) *J. Biol. Chem.* **263**, 12977–12985.
33. Schirmer, T., Bode, W. & Huber, R. (1987) *J. Mol. Biol.* **196**, 677–695.
34. Schoenleber, R. W., Lundell, D. J., Glazer, A. N. & Rapoport, H. (1984) *J. Biol. Chem.* **259**, 5481–5484.
35. Hsu, B.-D. (1988) *Plant Sci.* **55**, 199–203.
36. Gaboriaud, C., Bissery, V., Benchetrit, T. & Mornon, J. P. (1987) *FEBS Lett.* **224**, 149–155.
37. Ferguson, M. A. J. (1988) *Annu. Rev. Biochem.* **57**, 285–320.
38. Riethman, H. C., Mawhinney, T. P. & Sherman, L. A. (1988) *J. Bacteriol.* **170**, 2433–2440.
39. Kukuruzinska, M. A., Bergh, M. L. E. & Jackson, B. J. (1987) *Annu. Rev. Biochem.* **56**, 915–944.