JOURNAL OF BACTERIOLOGY, June 1987, p. 2675–2684
0021-9193/87/062675-10$02.00/0

Vol. 169, No. 6

Isolation and Characterization of Light-Regulated Phycobilisome Linker Polypeptide Genes and Their Transcription as a Polycistronic mRNA†

TERRI L. LOMAX,[1]‡ PAMELA B. CONLEY,[1] JAMES SCHILLING,[2] AND ARTHUR R. GROSSMAN[1]*

Department of Plant Biology, Carnegie Institution of Washington, Stanford, California 94305,[1] and California Biotechnology, Inc., Mountain View, California 94040[2]

Received 29 December 1986/Accepted 9 March 1987

Several cyanobacteria adjust both the phycobiliprotein and linker protein composition of the phycobilisome, a light-harvesting complex in cyanobacteria and some eucaryotic algae, to maximize absorption of prevalent wavelengths of light. This process is called complementary chromatic adaptation. We sequenced the amino terminus of a linker polypeptide which is associated with phycocyanin and accumulates to high levels during growth of the cyanobacterium *Fremyella diplosiphon* in red light. A mixed oligonucleotide encoding a region of this amino terminus was synthesized and used to identify a fragment of *F. diplosiphon* genomic DNA encoding the linker polypeptide. This linker gene was located between two other linker genes and contiguous to the red-light-induced phycocyanin gene set. Sequences of all three linker genes are presented. These genes were transcribed together onto a large polycistronic mRNA which also encoded the red-light-induced phycocyanin subunits. The relationship of this transcript to the biogenesis of the phycobilisome when *F. diplosiphon* is grown under different conditions of illumination is discussed.

PBSs are the major light-harvesting system of cyanobacteria (blue-green algae; procaryotes) and red algae (eucaryotes) (18–22, 49). They are water-soluble complexes attached to the outer surface of the photosynthetic membranes and may contain as many as 20 distinct polypeptides (18, 39). The PBS may provide 30 to 50% of the total light-harvesting capacity of the cells, absorbing light energy in the region of the visible spectrum where chlorophyll absorption is low and funneling this energy into the photosynthetic reaction centers (primarily photosystem II).

In the hemidiscoidal PBSs of cyanobacteria, there are two domains: the rods, which are constructed from hexameric stacks of the pigmented proteins (phycobiliproteins) PC and PE, and the core, which is composed of trimeric arrays of the phycobiliprotein AP. Each phycobiliprotein is composed of two related subunits, α and β, to which one or more tetrapyrrole chromophores are covalently attached via thioether linkages to cysteinyl residues. Together, the phycobiliproteins may account for 50% of the total soluble cell protein (4) and 85% of the mass of the PBS (45). The balance of the PBS (15%) is composed of linker polypeptides, which are essential for promoting assembly and maintaining the structure of the complex (20, 21, 32). In addition, linker polypeptides ensure efficient, unidirectional energy transfer from the peripheral pigments to the core and from the core to the chlorophyll of the reaction centers (23, 49). The linkers are basic proteins which associate tightly with the acidic biliproteins and range in molecular mass from 8 to 95 kDa. While most of the linker proteins are nonpigmented, some, such as the anchor protein (molecular weight between 75,000 and 95,000, depending on the organism), which is involved in the association of the PBS with the thylakoid membranes and energy transfer from the PBS to the reaction center of photosystem II (19, 32), contain a bilin chromophore.

There is only limited information about the amino acid composition and sequence of linker proteins and the genes encoding these proteins. Recently, either complete or partial sequences for linker polypeptides thought to be associated with the core and rod substructures of the PBS have been reported (16, 17). Füglistaller et al. (17) have shown that these linker polypeptides show homology to each other and also contain sequences homologous to the phycobiliproteins, suggesting that all of the genes encoding PBS components evolved from a single ancestral gene. The ancestral gene product appears to have some similarities with globin (41). Genes encoding the linker proteins have also recently been identified. The core linker gene is immediately downstream from the AP gene set in both *Anacystis nidulans* (27) and *Fremyella diplosiphon* (P. G. Lemaux and A. R. Grossman, manuscript in preparation) and is cotranscribed with the AP genes in the latter species (and probably in the former). Recently, linker genes have also been located downstream from the PC gene sets in *Synechococcus* sp. strain 7002 (7; D. A. Bryant, *in* T. Platt, ed., *Physiological Ecology of Picoplankton*, in press) and *Anabaena* sp. strain 7120 (2a).

In many cyanobacteria expression of both the linker proteins and the biliproteins is regulated by the wavelengths of light to which the algae are exposed. This ability to maximize absorption of prevalent wavelengths of light by changing the pigment composition of the PBS is called complementary chromatic adaptation. Growth in green light causes *F. diplosiphon* to synthesize high levels of PE while maintaining low levels of PC. In red light almost no PE is synthesized, while PC accumulates. Changes in the expression of PBS components during complementary chromatic adaptation appear to be mostly the result of differential

* Corresponding author.

† Carnegie Institution of Washington Plant Biology Publication no. 969.

‡ Present address: Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331.

transcription from phycobiliprotein genes (9, 35; A. Grossman, S. Robbins, P. Lemaux, and P. Conley, unpublished). In *F. diplosiphon* only a single PE gene set has been found (35; Grossman et al., unpublished data). Essentially no transcript encoding PE is detected in cells grown in red light, while an abundant transcript is present in cells grown in green light (35; Grossman et al., unpublished). In contrast, light-regulated PC synthesis involves differential expression from more than one gene set. In *F. diplosiphon* there are at least two gene sets encoding PC subunits (9, 10). One of these gene sets is constitutively expressed (transcripts from this gene set can be detected in both red and green light), while the other is only expressed in red light. Linker proteins also differentially accumulate during complementary chromatic adaptation (5), and it is important to determine how their levels are regulated relative to the phycobiliproteins with which they associate. In *F. diplosiphon* a minimum of five linker polypeptides have been associated with the rod substructure of the PBS (5). Two of these linkers accumulate coordinately with the red-light-induced PC subunits, two accumulate coordinately with the PE subunits, and one is constitutively expressed. In this manuscript we describe the isolation and characterization of the linker protein genes which are associated with the red-light-induced PC subunits and discuss the possible mechanisms involved in their regulated synthesis.

MATERIALS AND METHODS

Abbreviations. Abbreviations used in this paper include: kb, kilobase pair; AP, allophycocyanin; PC, phycocyanin; PE, phycoerythrin; PBS, phycobilisome; kDa, kilodalton; SDS, sodium dodecyl sulfate; bp, base pair; PAGE, polyacrylamide gel electrophoresis. Phycobiliprotein subunits and linker polypeptides of PBSs are abbreviated by the system of Glazer (19). The α and β subunits of a particular biliprotein, such as PC, are denoted PC^{α} and PC^{β}. Linker polypeptides are denoted L, and R, C, and M as subscripts to L indicate that the linker polypeptide is located in the rod, core, and membrane, respectively. Numerical superscripts to L give the apparent molecular masses (in kDa) of the linker polypeptides (as determined on SDS-polyacrylamide gels).

Materials. All chemicals were of reagent grade. Restriction enzymes were from Bethesda Research Laboratories and Boehringer Mannheim; calf intestine alkaline phosphatase, DNA polymerase I, and the Klenow fragment of DNA polymerase I were from Boehringer Mannheim; T4 DNA ligase was from Bethesda Research Laboratories; thio-[α-^{35}S]dATP (650 Ci/mmol; 1 Ci = 37 GBq), [γ-^{32}P]dATP (3,000 Ci/mmol), and [α-^{32}P]dCTP were from Amersham.

PBS isolation, protein electrophoresis, and purification. *F. diplosiphon*, a subculture of strain UTEX 481, was grown in Gorham medium (28) at 31°C in 5% CO_2–95% air and illuminated at 40 $\mu Em^{-2} s^{-1}$ (1 E = 1 mol of photons) with either red (Westinghouse F20T12/R) or green (Westinghouse F15T8/G) fluorescent lights. Intact PBSs were isolated from algal cells broken in a French pressure cell at 6,000 lb/in^2 and purified on sucrose step gradients (47). Electrophoretic separation of the polypeptides of isolated PBSs was done at a constant current of 10 mA on a 7.5 to 15% polyacrylamide gradient gel as described previously (13).

N-terminal microsequencing and synthesis of oligonucleotide probes. The linker polypeptide L_R^{39} was isolated by electroelution from polyacrylamide gels (31) and microsequenced with an Applied Biosystems 470A gas phase microsequencer (29, 37). The mixed oligonucleotide of 14 residues, encoding a portion of the amino-terminal sequence of L_R^{39}, was synthesized with the Applied Biosystems 388 oligonucleotide synthesizer by a modification of the standard phosphotriester method (43) and purified by PAGE (48).

Identification of linker genes on clones encoding phycobiliproteins. The mixed oligonucleotide probe was used to screen plaques of λ EMBL3 (15) previously demonstrated to encode biliprotein subunits (10). The oligonucleotide probe was radiolabeled (34) with [γ-^{32}P]dATP and T4 polynucleotide kinase. Plaque hybridizations (3) were performed at 37°C for 12 to 18 h in 0.5 M NaCl–0.1 M NaH_2PO_4–0.1 M Tris base–0.002 M EDTA–0.1% SDS (without prehybridization). Following hybridization, the nitrocellulose filters were washed four times (10 min each) at room temperature in 0.05 M PO_4 buffer, pH 7.0–0.002 M EDTA–0.1% SDS and exposed to Kodak XAR-5 film for 6 to 36 h at −80°C.

Restriction mapping of cloned sequences. Phage DNA from λ EMBL3 clone 37 (10) which hybridized to the oligonucleotide mixture was purified by the method of Maniatis et al. (34) and digested with the restriction enzymes *Pst*I, *Hin*dIII, and *Eco*RI as specified by the manufacturer. Fragments were resolved by electrophoresis in a 1.0% agarose gel in TBE buffer (0.05 M Tris, 0.05 M boric acid, 0.0025 M disodium EDTA) and transferred to nitrocellulose (34). Labeled oligonucleotide was hybridized to restriction fragments of clone 37 as described above for hybridization to plaque DNA.

DNA sequencing. DNA from clone 37 was digested with *Hin*dIII. A 1.8-kb fragment which hybridized to the oligonucleotide probe was purified from agarose gels by the method of Maniatis et al. (34) and either ligated directly into mp18 digested with *Hin*dIII or digested with *Pst*I and ligated into mp18 and mp19 digested with *Pst*I and *Pst*I-*Hin*dIII. *Sau*IIIA fragments of the isolated 1.8-kb fragment were also ligated into mp19 cut with *Bam*HI. Ligated DNA was used to transform *Escherichia coli* JM101 (36). Single-stranded DNA from plaques was isolated, and template was prepared and sequenced by the dideoxy chain termination method (36). Sequencing was also performed with oligonucleotide primers synthesized on the Biosearch 8600 oligonucleotide synthesizer. Primers were made to sequences approximately 200 nucleotides apart and used for dideoxy sequencing of the single-stranded template, containing the entire 1.8-kb *Hin*dIII insert.

A 3.8-kb *Hin*dIII fragment in plasmid pFD126 (9) adjacent (on the *F. diplosiphon* genome) to the 1.8-kb *Hin*dIII sequence encoding L_R^{39} was partially sequenced. pFD126 was digested with *Dde*I and electrophoresed on a 5% polyacrylamide gel in TBE buffer, and a 0.8-kb *Dde*I fragment was electroeluted from the gel in 0.2× TBE buffer (34). After the *Dde*I end was filled in with the Klenow fragment of DNA polymerase I, the 0.8-kb piece was digested with *Hin*dIII, yielding a 500-bp sequence contiguous with the 1.8-kb *Hin*dIII fragment. This piece was ligated into both mp18 and mp19 (digested with *Hin*dIII and *Hin*cII), and the recombinant molecules were used to transform JM101. The single-stranded template, in both orientations, was sequenced as described above. This fragment was also sequenced after the generation of progressive deletions with exonuclease III (25).

Comparison of linker amino acid sequences. Homologies among the linker polypeptides were examined with the protein sequence analysis program from Protein Identification Resources, Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.

Hybridization of linker gene sequences to RNA. RNA

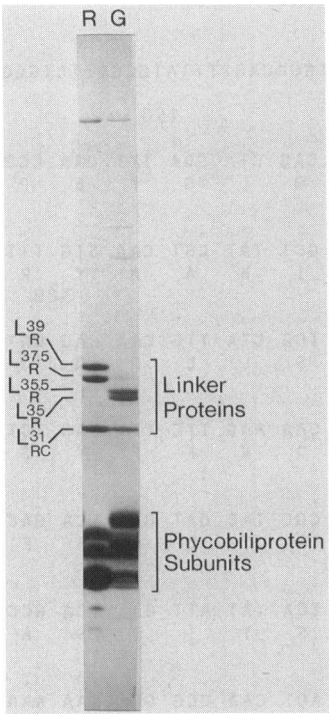

FIG. 1. SDS-PAGE of *F. diplosiphon* PBSs isolated from cells grown in red (R) or green (G) light. The name of each linker polypeptide, using the terminology of Glazer (17) (see Materials and Methods), is indicated on the left.

isolated from *F. diplosiphon* grown in both red and green light (9) was resolved by electrophoresis on a 1.5% agarose gel under denaturing conditions and transferred directly to nitrocellulose paper (46). Mixed oligonucleotide hexamers (Pharmacia) were used as primers to synthesize radiolabeled probes (14) from either single-stranded M13 templates or heat-denatured, double-stranded DNA fragments. Hybridization of the radiolabeled DNA to the RNA and conditions for washing the nitrocellulose filters following hybridization have been described (9).

S1 nuclease mapping. The 3′ end of the large (3,800-base) transcript encoding both the red-light-induced PC subunits and linker polypeptides was determined by S1 nuclease protection (8). A 291-bp *Hin*dIII-*Pst*I fragment encoding the 3′ end of $L^{9.7}_{R}$ (plus additional downstream sequences) was cloned into mp19, radiolabeled by complementary second-strand synthesis with 1 μg of template, 1.5 U of the Klenow fragment of DNA polymerase I, and the 17-base M13 sequencing primer (Pharmacia, no. 27-1534-01). Unincorporated nucleotides were separated from the labeled complementary strand by column chromatography on a Sephadex G-50 column. The reaction product was cleaved with *Pst*I for 1 h at 37°C (generating a labeled 301-base fragment), and one-fifth of the labeled DNA was precipitated together with 26.5 μg of RNA from red- or green-light-grown cultures. RNA-DNA hybridization (39°C) and S1 digestions were performed as described previously (30). Following digestion, the nucleic acid was prepared as for sequencing and loaded on a 6% sequencing gel. The size of the protected fragment was determined by comparison with a sequencing ladder generated for mp18 from the 17-base sequencing primer. Exposure to Kodak XAR-5 film was done for 24 h at −80°C with an intensifying screen.

RESULTS

PBSs can be readily isolated as intact light-harvesting complexes (18). This allowed us to compare the polypeptide composition of PBSs from *F. diplosiphon* grown under various conditions of illumination. PBSs from *F. diplosiphon* contain phycobiliprotein subunits ($PC^{\alpha,\beta}$, $AP^{\alpha,\beta}$, and $PE^{\alpha,\beta}$) and linker polypeptides (L^{31}_{RC}, L^{35}_{R}, $L^{35.5}_{R}$, $L^{37.5}_{R}$, and L^{39}_{R}) which exhibit differential accumulation during growth in red and green light (Fig. 1) (4–6). The sizes of the linker proteins (39.0, 37.5, 35.5, 35.0, and 31.0 kDa) were estimated from migration on denaturing polyacrylamide gels. Linker L^{31}_{RC} was present during growth of *F. diplosiphon* in both red and green illumination (its level did not appear to change significantly with different light qualities) and was therefore likely to be involved in the association of the core of the PBS with the first PC hexamer of the rod substructure. The linker polypeptides L^{39}_{R} and $L^{37.5}_{R}$ were present in PBSs of cells grown in red light and either absent or nearly absent in cells grown in green light, suggesting that these linkers associate with the PC subunits synthesized during growth of *F. diplosiphon* in red light. Some low-molecular-weight polypeptides were also present in PBSs from cells cultured in red light but not from green-light-grown cells (below the phycobiliprotein subunits in Fig. 1). The proteins $L^{35.5}_{R}$ and L^{35}_{R}, often represented by a doublet on polyacrylamide gels, were present only in PBSs from green-light-grown *F. diplosiphon*, as were the PE subunits (Fig. 1), suggesting that these linkers are associated with the red-pigmented PE hexamers (involved in interactions of PE-PC and PE-PE hexamers).

To demonstrate the presence of the gene encoding the linker polypeptide L^{39}_{R} on the clones containing fragments of *F. diplosiphon* genomic DNA, we constructed a 14-base mixed oligonucleotide encoding a region of the amino termi-

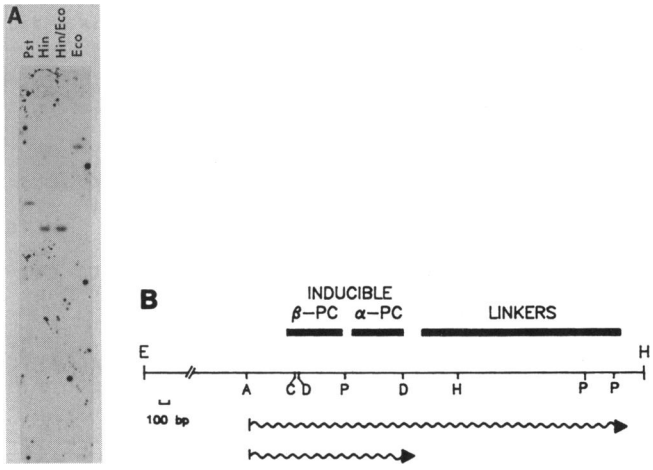

FIG. 2. Location of the gene encoding L^{39}_{R} by oligonucleotide hybridization to restriction fragments of λ clone 37 DNA. (A) Hybridization of oligonucleotide to clone 37 DNA digested with *Pst*I, *Hin*dIII, *Hin*dIII-*Eco*RI, and *Eco*RI. The sizes of each hybridizing species were; *Pst*I, 2.4 kb; *Hin*dIII, 1.8 kb; *Hin*dIII-*Eco*RI, 1.8 kb; and *Eco*RI, 6.0 kb. Markers used to estimate the sizes of the hybridizing fragments were generated by digestion of λ DNA with *Sal*I, *Hin*dIII, or *Eco*RI. (B) Map of the region of the *F. diplosiphon* genome which hybridized with the oligonucleotide (marked linkers) and contained the inducible PC gene set. H, P, E, C, A, and D, Restriction enzyme sites for *Hin*dIII, *Pst*I, *Eco*RI, *Cla*I, *Aha*III, and *Dde*I, respectively. The wavy arrows indicate transcripts generated from this region of the genome.

```
                                                  50
AGC TAG GGTTTCCTTTCGCCCAGAACAACTAAAGGCTGTGGATGCTAATTTCTAGCATGGGCAGTTTTATGGGGTTCTGGGCATGTTTTTTCT
 S
    100                                                                       150
GATTGAAATTCAGGAGATTCAGTC    ATG ACT AGT TCC ACA GCG GCC CGA CAG CTA GGA TTT GAA CCG TTT GCT AGC
                             M   T   S   S   T   A   A   R   Q   L   G   F   E   P   F   A   S
                                        200
ACA GCT CCC ACA GAA TTG AGA GCT AGT AGT GAC GTA ATT CAT GCT GCT TAT CGT CAA GTG TTT CAA GTG TTT
 T   A   P   T   E   L   R   A   S   S   D   V   P   A   V   I   H   A   A   Y   R   Q   V   F
        250                                                                 300
GGC AAT GAC CAC GTG ATG CAA AGC GAA CGT CTC ACC AGT GCA GAA TCG CTA TTG CAA CAG GGT AAC ATT AGC
 G   N   D   H   V   M   Q   S   E   R   L   T   S   A   E   S   L   L   Q   Q   G   N   I   S
                                                350
GTC AGA GAT TTT GTG CGG TTA TTA GCA CAA TCT GAA CTA TAT CGC CAA AAG TTC TTT TAC TCT ACA CCT CAA
 V   R   D   F   V   R   L   L   A   Q   S   E   L   Y   R   Q   K   F   F   Y   S   T   P   Q
                    400                                                             450
GTC CGC TTC ATC GAA CTG AAC TAC AAG CAT TTG CTT GGT CGT GCG CCC TAC GAT GAA TCA GAG ATT TCT TAC
 V   R   F   I   E   L   N   Y   K   H   L   L   G   R   A   P   Y   D   E   S   E   I   S   Y
                                                            500
CAC GTC AAC CTT TAT ACA GAG AAA GGA TAC GAG GCA GAA ATC AAT TCA TAT ATT GAT TCA GCC GAA TAT CAA
 H   V   N   L   Y   T   E   K   G   Y   E   A   E   I   N   S   Y   I   D   S   A   E   Y   Q
                                550                                                         600
GAA AGC TTT GGC GAA CGA ATT GTA CCT CAT TAT CGG GGA TTT GAA ACT CAA CCG GGA CAA AAG ACA GTA GGC
 E   S   F   G   E   R   I   V   P   H   Y   R   G   F   E   T   Q   P   G   Q   K   T   V   G
                                                                        650
TTC AAC CGG ATG TTT CAG ATT TAT CGG GGA TAT GCT AAC AGC GAT CGC TCC CAA GGT AAA AAC AAG TCT GCT
 F   N   R   M   F   Q   I   Y   R   G   Y   A   N   S   D   R   S   Q   G   K   N   K   S   A
                                    700
TGG TTA ACC CAA GAT CTA GCA CTG AAT TTG GCT AGC AAC ATT CAG ACT CCT AAC TTT GGC AAG GGA CTG ACT
 W   L   T   Q   D   L   A   L   N   L   A   S   N   I   Q   T   P   N   F   G   K   G   L   T
    750                                                                   800
GGT GTA GTC GCA GGC GAT CGC GGA CAG CTT TAC CGA GTA CGA GTC ATC CAA GCA GAT AGA GGG CGT ACT ACT
 G   V   V   A   G   D   R   G   Q   L   Y   R   V   R   V   I   Q   A   D   R   G   R   T   T
                                                850
CAA ATC AGA CGT AGT ATT CAA GAG TAT TTG GTC AGT TAC GAC CAA CTT TCA CCA ACT CTG CAA AGA CTG AAT
 Q   I   R   R   S   I   Q   E   Y   L   V   S   Y   D   Q   L   S   P   T   L   Q   R   L   N
            900                                                             950
CAG CGA GGT AGT CGC GTA GTT AAT ATT TCT CCT GCC TAA ACTGAATCAAATTTGCCACATCAAAAAGTCTACAAGGAGAGCT
 Q   R   G   S   R   V   V   N   I   S   P   A END
                                            1000
ATTCGTG     CCA ATT ACA TCC GCA GCA TCT CGT TTG GGA ACT ACA GCT TAC CAA ACA AAT CCT ATT GAA CTG
             P   I   T   S   A   A   S   R   L   G   T   T   A   Y   Q   T   N   P   I   E   L
            1050                                                            1100
CGT CCA AAC TGG ACA GCA GAA GAT GCC AAG ATT GTT ATC CAG GCA GTC TAT CGT CAG GTA TTA GGC AAT GAC
 R   P   N   W   T   A   E   D   A   K   I   V   I   Q   A   V   Y   R   Q   V   L   G   N   D
                                                1150
TAC TTG ATG CAG TCA GAA CGT CTC ACC AGC TTA GAG TCG TTG TTG ACC AAT GGT AAA CTG AGC GTC AGA GAT
 Y   L   M   Q   S   E   R   L   T   S   L   E   S   L   L   T   N   G   K   L   S   V   R   D
                    1200                                                            1250
TTT GTC CGC GCA GTA GCT AAA TCA GAA TTA TAC AAA ACA AAG TTT CTT TAC CCA CAT TTT CAA ACT CGG GTA
 F   V   R   A   V   A   K   S   E   L   Y   K   T   K   F   L   Y   P   H   F   Q   T   R   V
                                                            1300
ATT GAA TTG AAC TTT AAG CAT TTG TTA GGA AGA GCT CCT TAC GAT GAA TCG GAA GTG ATT GAG CAT CTC GAT
 I   E   L   N   F   K   H   L   L   G   R   A   P   Y   D   E   S   E   V   I   E   H   L   D
                                1350
CGC GAC CAA AAT CAA GGG TTC GAC GCA GAC ATC GAT TCC TAT ATA GAT TCT GCT GAA TAT GAT ACC TAT TTT
 R   Y   Q   N   Q   G   F   D   A   D   I   D   S   Y   I   D   S   A   E   Y   D   T   Y   F
1400                                                                1450
```

FIG. 3. Nucleotide sequence encoding linker polypeptides. A DNA sequence of 2,369 nucleotides encoding the final amino acid of PC^{α} (serine) and three proteins, $L_R^{37.5}$, L_R^{39}, and $L_R^{9.7}$, is presented. Ribosome-binding sites preceding each of the coding regions are underlined. The initiator codon for L_R^{39} (positions 975 to 977) is also underlined. The sequence is numbered at 50-nucleotide intervals, with dots above every 10th nucleotide.

nus of L_R^{39}. The first 20 amino acids of L_R^{39}, obtained by microsequencing the isolated polypeptide, were PITSA ASRLGTTAYQTNPIE. An oligonucleotide, TA(C/T)CA (G/A)AC(G/A/C/T)AA(C/T)CC, which encodes amino acids 14 to 18 (excluding the last base of the proline codon), was synthesized. Since many genes encoding PBS components appear to be clustered on cyanobacterial genomes (reviewed in reference 24), labeled oligonucleotide was used to screen λ EMBL3 clones encoding phycobiliprotein subunits. Clone 37, a previously characterized clone (10) containing genes for two sets of PC and one set of AP subunits, hybridized to this oligonucleotide. The oligonucleotide hybridized to a 1.8-kb *Hin*dIII fragment, a 2.4-kb *Pst*I fragment, and a 6.0-kb *Eco*RI fragment from clone 37. *Eco*RI did not cut the 1.8-kb *Hin*dIII fragment (as shown in the *Hin*dIII-*Eco*RI digest in Fig. 2). With a previously developed restriction map (10), the region of the clone homologous to the oligonucleotide sequence was located just downstream from the red-light-induced PC subunit genes, a region designated "linkers" in Fig. 2.

Sequence analyses of the 1.8-kb *Hin*dIII fragment plus a portion of the preceding *Hin*dIII fragment which encodes PC^{α} and PC^{β} (Fig. 2, bottom) revealed three open reading frames contiguous to the PC gene set, which were designated *lpcA*, *lpcB*, and *lpcC*. In Fig. 3 the complete nucleotide sequence from the last amino acid (serine) encoded by the PC^{α} gene to the 3′ end of the 1.8-kb *Hin*dIII fragment is presented, with the deduced amino acid sequences of the open reading frames shown directly below the nucleotide sequence. The amino-terminal amino acid sequence encoded by the second open reading frame (*lpcB*, nucleotides 978 to 1841, Fig. 3) exactly matched the sequence derived by microsequencing the amino terminus of L_R^{39}. The first open reading frame (*lpcA*), which was separated from the PC^{α} gene by 114 nucleotides and from *lpcB* by 53 nucleotides, encoded a protein highly homologous to L_R^{39} (Fig. 4A) and

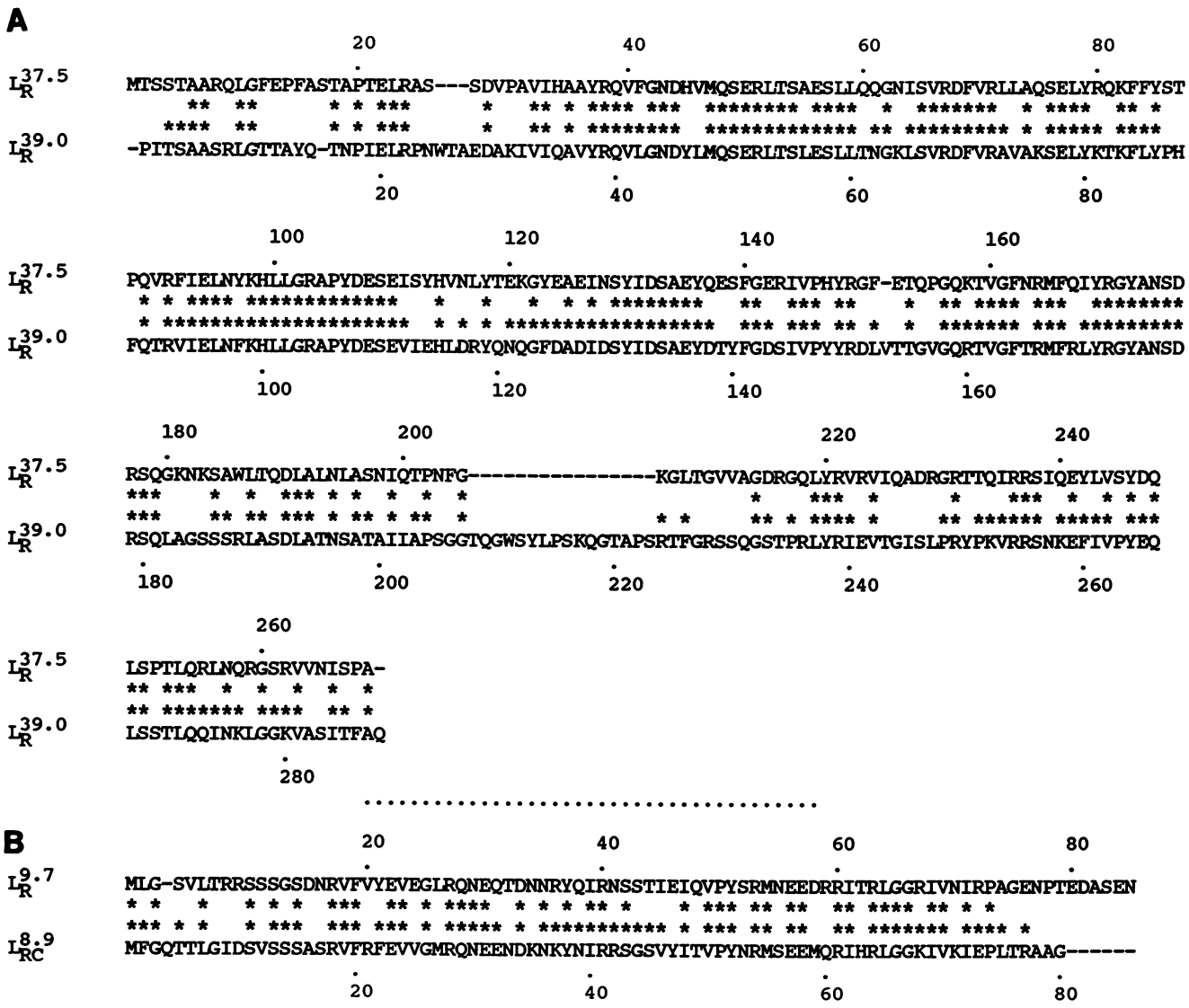

FIG. 4. (A) Comparison of amino acid sequence, as deduced from nucleotide sequence, of the first open reading frame (nucleotides 118 to 924, polypeptide designated $L_R^{37.5}$) following the red-light-induced PC gene set with L_R^{39} (encoded by nucleotides 978 to 1841). (B) Comparison of amino acid sequence, as deduced from the nucleotide sequence, of the last open reading frame (nucleotides 1860 to 2114, polypeptide designated $L_R^{9.7}$) with $L_{RC}^{8.9}$ of *M. laminosus* (17). Identical amino acids are indicated by two stars, and conserved amino acid changes by a single star (1).

was designated $L_R^{37.5}$. It is very likely that the protein encoded by *lpcA* is the linker protein $L_R^{37.5}$ shown in Fig. 1, because it was highly homologous to but slightly smaller than L_R^{39} (269 versus 288 amino acids) and the transcript from the gene was only present in *F. diplosiphon* grown in red light (see below), as was $L_R^{37.5}$.

Overall homology of $L_R^{37.5}$ and L_R^{39} (Fig. 4A) was 53%, and specific regions of the two proteins were identical or nearly identical. For example, amino acids 28 to 71 of $L_R^{37.5}$ were 70% homologous to amino acids 29 to 72 of L_R^{39}, and amino acids 91 to 110 of $L_R^{37.5}$ were 90% homologous to amino acids 92 to 111 of L_R^{39}. Furthermore, many positions had conserved amino acid substitutions. Both $L_R^{37.5}$ and L_R^{39} had pIs of 10.55. The carboxyl ends of the proteins were especially basic, with a pI of 12.32 for amino acids 166 to 269 of $L_R^{37.5}$ and 12.44 for amino acids 168 to 288 of L_R^{39}. Clustering of positively charged amino acids was even more apparent from position 220 to the end of each protein. This basic region of each protein may be important in interacting with the acidic groups that are exposed in the hexameric arrays of the phycobiliproteins (49). As deduced from Fig. 4A, the difference in size between $L_R^{37.5}$ and L_R^{39} could be mostly accounted for by a small insertion preceding the region of the gene encoding amino acid 206 of $L_R^{37.5}$ (homology between the two proteins broke down after amino acid 207 of L_R^{39} and was reestablished at amino acid 224), which suggests that the genes encoding these proteins diverged from each other via a small DNA insertion of deletion.

A third open reading frame (*lpcC*) of 85 amino acids, which began 18 nucleotides downstream from the end of *lpcB* (Fig. 3), encoded a polypeptide which was highly homologous (47% overall) to a polypeptide, $L_{RC}^{8.9}$ (Fig. 4B), sequenced by Füglistaller et al. (17) and thought to be associated with PC of the rods of the PBS, and to a 9-kDa rod linker from *Synechococcus* sp. strain 7002 (7; D. Bryant, personal communication). It was also homologous to the

TABLE 1. Codon usage for *F. diplosiphon* linker genes

Codon	Amino acid	No. of occurrences			Codon	Amino acid	No. of occurrences		
		$L_R^{37.5}$	L_R^{39}	$L_R^{9.7}$			$L_R^{37.5}$	L_R^{39}	$L_R^{9.7}$
UUU	F	9	8	1	UCU	S	5	10	2
UUC	F	3	3	0	UCC	S	2	4	1
UUA	L	3	6	2	UCA	S	4	2	2
UUG	L	5	6	2	UCG	S	1	3	0
CUU	L	4	4	0	CCU	P	4	4	2
CUC	L	1	2	0	CCC	P	2	3	0
CUA	L	4	3	0	CCA	P	2	6	1
CUG	L	5	5	0	CCG	P	2	0	0
AUU	I	9	8	4	ACU	U	7	7	2
AUC	I	4	6	2	ACC	U	2	7	2
AUA	I	0	2	0	ACA	U	6	9	0
AUG	M	3	2	2	ACG	U	0	0	1
GUU	V	1	6	2	GCU	A	8	6	0
GUC	V	6	5	2	GCC	A	3	1	0
GUA	V	7	4	2	GCA	A	7	12	2
GUG	V	3	1	0	GCG	A	2	2	0
UAU	Y	9	7	1	UGU	C	0	0	0
UAC	Y	7	11	2	UGC	C	0	0	0
UAA	—	0	0	0	UGA	—	0	0	0
UAG	—	0	0	0	UGG	W	1	2	0
CAU	H	3	3	0	CGU	R	5	9	3
CAC	H	2	0	0	CGC	R	5	6	5
CAA	Q	16	10	2	CGA	R	5	0	0
CAG	Q	6	6	2	CGG	R	4	2	1
AAU	N	5	7	5	AGU	S	7	3	1
AAC	N	8	3	4	AGC	S	6	8	3
AAA	K	2	6	0	AGA	R	5	6	3
AAG	K	5	5	0	AGG	R	0	0	0
GAU	D	7	9	3	GGU	G	5	14	3
GAC	D	3	5	1	GGC	G	5	1	2
GAA	E	12	10	7	GGA	G	7	4	1
GAG	E	4	3	2	GGG	G	1	1	0

small linker found in the core of the PBS (27). Therefore, it is likely that this open reading frame encodes a linker polypeptide. The molecular size of this protein was 9.7 kDa, as deduced from the sequence of the open reading frame, and it was designated $L_R^{9.7}$. Since this small PBS component was translated from a red-light-induced RNA (see below), it may be associated with the red-light-induced PC hexamers of the rods.

All three of the open reading frames presented in Fig. 3 were preceded by a ribosome-binding site (AGGAGA) identical to the ribosome-binding sites preceding the initiator codons on the transcripts for PE^{α} and PE^{β} (Grossman et al., unpublished) and similar or identical to ribosome-binding sites on the constitutive PC (10; Conley and Grossman, unpublished data) and AP (Lemaux and Grossman, unpublished data) mRNAs. While translation of $L_R^{37.5}$ and $L_R^{9.7}$ RNA was initiated with a methionine codon, no such codon was present at the beginning of the coding region for L_R^{39}, and the mature polypeptide began with a proline. Therefore, translation initiation may begin at GTG at position 975 (Fig. 3).

Table 1 shows codon usage for the linker genes. Comparison of codon usage in the linker genes and biliprotein genes (AP, PE, and PC) of *F. diplosiphon* (35; Lemaux and Grossman, manuscript in preparation; Conley and Grossman, manuscript in preparation) revealed some differences. In general there was less codon bias in the sequences of the linker genes. The codon AGG (arginine) was not found in any of the biliprotein genes sequenced in *F. diplosiphon* and was not present in the three linker genes. However, the codons GUG (Valine), CCG (proline) and UCG (serine), which were not present in PC (Conley and Grossman, manuscript in preparation), AP (Lemaux and Grossman, manuscript in preparation), or PE (36, Grossman et al., unpublished) subunit genes, were used to a limited extent in *lpcA* and *lpcB* (not in *lpcC*).

The cluster of linker genes hybridized to a red-light-induced transcript of approximately 3,800 bases (Fig. 5). This transcript, along with a more prominent 1,600-base red-light-induced transcript, hybridized with a DNA fragment encoding a specific PC gene set (9) (Fig. 5, lanes 1 and 2). The 1.8-kb *Hin*dIII fragment containing the 3′ end of *lpcA* and the entire *lpcB* and *lpcC* genes hybridized only to the red-light-induced 3,800-base transcript (Fig. 5, lanes 3 and 4). A *Pst*I fragment containing the end of *lpcB* and most of the *lpcC* coding region and a *Pst*I-*Hin*dIII fragment containing the end of *lpcC* (see the legend to Fig. 5 for details) also hybridized to the 3,800-base transcript (Fig. 5, lanes 5 to 8). Structural analyses of the region of the transcript which lies between the sequences encoding PC^{α} and $L_R^{37.5}$ (*lpcA*) (Conley and Grossman, manuscript in preparation) and the

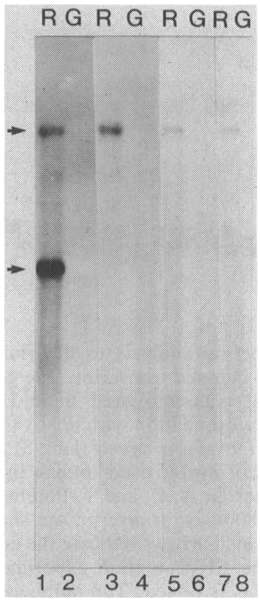

FIG. 5. Hybridization of DNA fragments encoding PC and linker proteins to RNA from *F. diplosiphon*. Lanes 1 and 2, Hybridization of a 900-bp *Dde*I fragment from pFD126 (9) encoding the red-light-induced PC^{α} and PC^{β} to RNA from cells grown in red (R) or green (G) light. Lanes 3 and 4, Hybridization of the 1.8-kb *Hin*dIII fragment containing the last 393 nucleotides of *lpcA* and the complete sequences for *lpcB* and *lpcC* (nucleotides 531 to 2369 of Fig. 3) to RNA from red- and green-light-grown cells. Lanes 5 and 6, Hybridization of a *Pst*I fragment containing the last 49 nucleotides of *lpcB* and the first 222 nucleotides of *lpcC* (nucleotides 1794 to 2081 of Fig. 3) to RNA from red- and green-light-grown cells. Lanes 7 and 8, Hybridization of the *Hin*dIII-*Pst*I fragment containing the last 33 nucleotides of *lpcC* (nucleotides 2082 to 2369 of Fig. 3) to RNA from red- and green-light-grown cells. Arrows indicate the positions of the 3,800- and 1,600-base transcripts (upper and lower arrows, respectively).

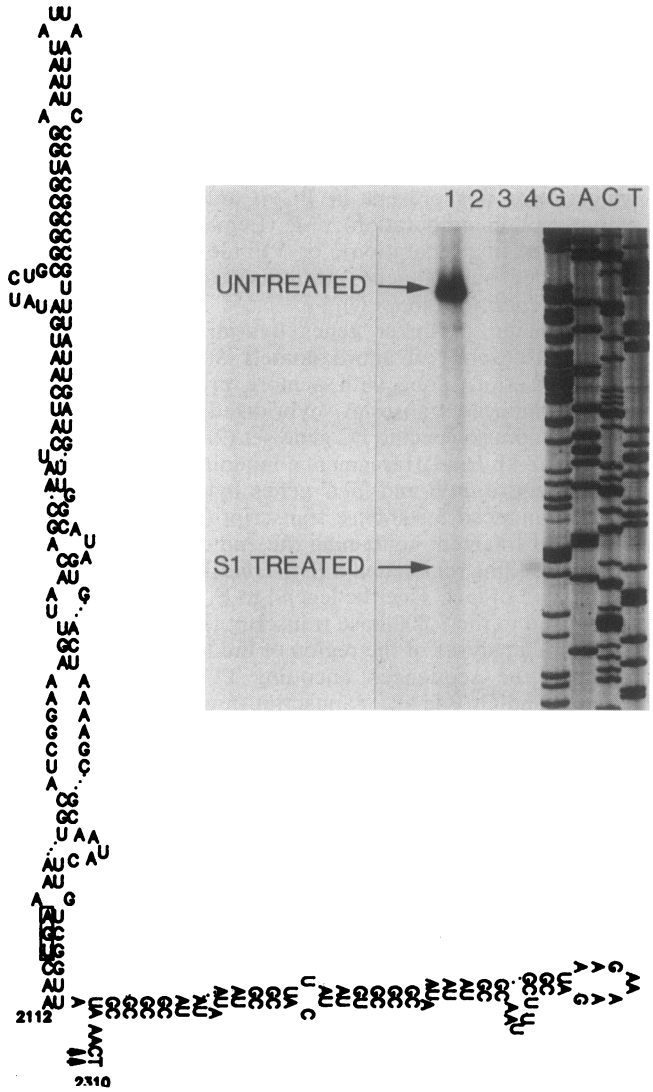

FIG. 6. S1 protection analysis to determine the 3′ end of the 3,800-base red-light-induced transcript. The gel shows a ^{32}P-labeled DNA fragment that was untreated (lane 1) or treated with S1 nuclease following hybridization with tRNA (lane 2), and total RNA from *F. diplosiphon* grown in green (lane 3) and red (lane 4) light. Sequencing ladders of mp18, used to size the protected fragment, are presented in lanes G, A, C, and T. Potentially stable hairpins at the 3′ end of the 3,800-base transcript are shown to the left of and below the gel. The small arrows indicate the end of the transcript, as determined from the S1 protection experiments. The termination codon (UGA) of *lpcC* is boxed. The numbers at the beginning and end of the sequence refer to the number of the nucleotide as presented in Fig. 3.

sequence immediately following the region encoding $L_{RC}^{9.7}$ (*lpcC*) (Fig. 6) showed a strong potential for secondary structure. S1 mapping of the 3′ end of the 3,800-base transcript is presented in Fig. 6. No protection of the DNA fragment (lane 1) from S1 nuclease was observed after it was incubated with either tRNA (lane 2) or total RNA from green-light-grown *F. diplosiphon* (lane 3). RNA from red-light-grown *F. diplosiphon* protected a 232- to 233-base DNA fragment (the size was determined by comparison with sequence ladders of mp18, lanes G, A, C, T), locating the 3′ end of the large red-light-induced transcript at nucleotides 2309 to 2310 (Fig. 3). The last nucleotide of the transcript was adjacent to a potential hairpin structure ($\Delta G = -93$ kcal for the entire secondary structure shown in Fig. 6), which may serve in transcription termination.

DISCUSSION

In this manuscript we report the complete sequence of three genes, designated *lpcA*, *lpcB*, and *lpcC*, which encode linker polypeptides and establish that these genes are transcribed into a 3,800-base red-light-induced mRNA which also encodes PC subunits. Identification of the gene for the highest-molecular-weight linker was established because the amino-terminal sequence of L_R^{39} and the sequence encoded by the 5′ end of *lpcB* were identical. The open reading frame immediately upstream (gene designation *lpcA*) from L_R^{39} and downstream from the PC^{α} gene also encoded a linker, based on its considerable homology with L_R^{39}. Since this open reading frame contained information for a protein slightly smaller than L_R^{39} (269 versus 288 amino acids) and was induced in red light, it most likely encodes the red-light-induced linker designated $L_R^{37.5}$ in the gel presented in Fig. 1. The molecular sizes determined for L_R^{39} and $L_R^{37.5}$ from the sequences of *lpcA* and *lpcB* were smaller (32.5 versus 39 kDa for L_R^{39} and 30.5 versus 37.5 kDa for $L_R^{37.5}$) than the apparent sizes determined by gel electrophoresis. Molecular weights predicted from migration on denaturing polyacrylamide gels are frequently unreliable (11, 26, 40), and it is possible that the anomolous migration of these polypeptides is the result of their extremely basic isoelectric points. Finally, the last open reading frame was identified as a PBS component by its strong homology to $L_{RC}^{8.9}$ of *Mastigocladus laminosus*, which was sequenced at the protein level by Füglistaller et al. (17), and to a 9-kDa polypeptide (capping linker) involved in the association of the terminal PC trimer with the rod substructure (7, Bryant, personal communication). Because it is present on the red-light-induced transcript, it is probably associated with the distal, red-light-induced PC hexamers and may represent the capping linker of the *F. diplosiphon* PBS.

Changes in PBS components during adaptation to different wavelengths of illumination appear to be controlled, to a good extent, at the level of transcription. The levels of the two transcripts (1,600 and 3,800 bases) which encode the peripheral rod components present in red-light-grown cells may reflect the ratios between phycobiliprotein subunits and the linker polypeptides. However, we have determined, by densitometer tracings of Northern blots (data not shown), that there is a six- to sevenfold difference in the levels of the 1,600- and 3,800-base transcripts. This would mean that there are between seven and eight sets of sequences encoding the PC subunits for every two sequences encoding a linker polypeptide. Since approximately six sets of PC subunits should be present for each of the two linkers (e.g., one might expect 10 to 12 copies of the 1,600-base transcript for each 3,800-base transcript if the level of control were solely transcriptional), the relative abundance of the two different-sized transcripts may not be the only factor governing the final levels of the protein product. Codon usage among the different genes may be another factor which results in differential accumulation of PBS polypeptides. Some codons used in the translation of linker RNA sequences are different from those used in the PC (Conley and Grossman, manuscript in preparation), AP (Lemaux and Grossman, manuscript in preparation), and PE transcripts (35; Grossman et al., unpublished). If these codons are not

matched by prevalent tRNA species, it is possible that translational efficiency also plays a role in maintaining the different stoichiometries of the PBS components. In *E. coli* the use of rare codons can cause decreased translation of mRNAs and has been associated with less abundantly expressed proteins (12). Furthermore, the RNA encoding L_R^{39} does not start with the initiator codon AUG, but may start with GUG (indicated as GTG in the DNA sequence presented). While GUG and UUG can be used as an initiator codon in *E. coli* a small percentage of the time (44) (approximately 8% for GUG and 1% for UUG), generally these codons are less efficient in allowing translation initiation than AUG (42, 44). Since the L_R^{39} polypeptide begins with a proline, as determined by the amino-terminal sequence, the initiator methionine must be removed either prior to or during formation of the PBS.

It is still uncertain what factors govern the ratio of the 1,600- and 3,800-base transcripts. The smaller transcript may be the consequence of termination of transcription or may result from processing of the large transcript and degradation of the segment encoding the linkers to yield the more stable 1,600-base transcript. A strong potential for secondary structure was observed at the end of the large transcript (Fig. 6) and also between the regions of the transcript encoding PC^{α} and $L_R^{37.5}$ (not shown), but it is still uncertain whether these structures play a role in transcription termination or are sites for RNA processing. Segmental stability of RNA transcripts has been invoked as the cause of differences in the abundance of a large 2.7-kilobase and small 0.59- and 0.49-kilobase transcripts from *Rhodobacter capsulatus* (2). In this case the small transcript encoding the light-harvesting polypeptides is generated via degradation of the region of the large transcript encoding the reaction center polypeptides, and the different levels of specific RNA species may completely account for the different levels of the light-harvesting and reaction center polypeptides. Examination of the relationship between the large and small PC-encoding transcripts may enable us to learn about factors which contribute to transcription termination, RNA stability and processing, and the regulated accumulation of the different PBS polypeptides in cyanobacteria.

ACKNOWLEDGMENTS

We acknowledge Sabrina Robbins, Loretta Tayabas, and Glenn Ford for excellent technical assistance. We also thank Peggy Lemaux and Lamont Anderson for helpful discussions and Donald Bryant and William Belknap for sharing their unpublished data with us.

T.L.L. was supported by a plant molecular biology postdoctoral fellowship from the National Science Foundation (DMD-8312552). P.B.C. was supported by a Carnegie Institution of Washington Fellowship. A.R.G. was supported by the Carnegie Institution of Washington and Public Health Service grant GM-33436-01 from the National Institutes of Health.

LITERATURE CITED

1. **Bacon, D. J., and W. F. Anderson.** 1985. Multiple sequence alignment. J. Mol. Biol. **191**:153–161.
2. **Belasco, J. G., T. Beatty, C. W. Adams, A. von Gabain, and S. N. Cohen.** 1985. Differential expression of photosynthesis genes in *Rhodopseudomonas capsulata* results from segmental differences in stability within the polycistronc *rxc* transcript. Cell **40**:171–181.

2a. **Belknap, W. R., and R. Haselkorn.** 1987. Cloning and light regulation of expression of the phycocyanin operon of the cyanobacterium *Anabaena*. EMBO J. **6**:871–884.

3. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
4. **Bogorad, L.** 1975. Phycobiliproteins and complementary chromatic adaptation. Annu. Rev. Plant Physiol. **26**:369–401.
5. **Bryant, D. A.** 1981. The photoregulated expression of multiple phycocyanin species. General mechanism for control of phycocyanin synthesis in chromatically adapting cyanobacteria. Eur. J. Biochem. **119**:425–429.
6. **Bryant, D. A., and G. Cohen-Bazire.** 1981. Effects of chromatic illumination on cyanobacterial phycobilisomes. Evidence for the specific induction of a second pair of phycocyanin subunits in *Pseudanabaena* 7409 grown in red light. Eur. J. Biochem. **119**:415–424.
7. **Bryant, D. A., R. De Lorimier, G. Guglielmi, V. L. Stirewalt, A. Cantrell, and S. E. Stevens, Jr.** 1986. The cyanobacterial photosynthetic apparatus, p. 749–755. *In* J. Biggin (ed.), Progress in photosynthesis (Proceedings of the VIIth International Congress on Photosynthesis). Martinus Nijhoff Publishers, The Hague.
8. **Clary, D. O., and D. R. Wolstenholme.** 1985. The ribosomal RNA genes of Drosophila mitochondrial DNA. Nucleic Acids Res. **13**:4029–4045.
9. **Conley, P. B., P. G. Lemaux, and A. R. Grossman.** 1985. Cyanobacterial light-harvesting complex subunits encoded in two red light-induced transcripts. Science **230**:550–553.
10. **Conley, P. B., P. G. Lemaux, T. L. Lomax, and A. R. Grossman.** 1986. Genes encoding major light-harvesting polypeptides are clustered on the genome of the cyanobacterium *Fremyella diplosiphon*. Proc. Natl. Acad. Sci. USA **83**:3924–3928.
11. **Darley-Usmar, M. V., and S. D. Fuller.** 1981. M_r-values of mature subunits I and III of beef heart cytochrome *c* oxidase in relationship to nucleotide squences of their genes. FEBS Lett. **135**:164–166.
12. **de Boer, H. A., and R. A. Kastelein.** 1986. Biased codon usage: an explanation of its role in optimization of translation, p. 225–285. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, London.
13. **Egelhoff, T., and A. R. Grossman.** 1983. Cytoplasmic and chloroplast synthesis of phycobilisome polypeptides. Proc. Natl. Acad. Sci. USA **80**:3339–3343.
14. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
15. **Frischauf, A. M., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.
16. **Füglistaller, P., R. Rumbeli, F. Suter, and H. Zuber.** 1984. Minor polypeptides from the phycobilisome of the cyanobacterium *Mastigocladus laminosus*. Hoppe-Seyler's Z. Physiol. Chem. **365**:1085–1096.
17. **Füglistaller, P., F. Suter, and H. Zuber.** 1985. Linker polypeptides of the phycobilisome from the cyanobacterium *Mastigocladus laminosus*: amino acid sequence and relationships. Biol. Chem. Hoppe-Seyler **366**:993–1001.
18. **Gantt, E.** 1981. Phycobilisomes. Annu. Rev. Plant Physiol. **32**:327–347.
19. **Glazer, A. N.** 1985. Light harvesting by phycobilisomes. Annu. Rev. Biochem. Biophys. **14**:47–77.
20. **Glazer, A. N.** 1984. A macromolecular complex optimized for light energy transfer. Biochim. Biophys. Acta **768**:29–51.
21. **Glazer, A. N.** 1982. Phycobilisomes: structure and dynamics. Annu. Rev. Microbiol. **36**:173–198.
22. **Glazer, A. N., D. J. Lundell, G. Yamanaka, and R. C. Williams.** 1983. The structure of a simple phycobilisome. Ann. Microbiol. Inst. Pasteur (Paris) **134B**:159–180.
23. **Glazer, A. N., S. W. Yeh, S. P. Webb, and J. H. Clark.** 1985. Disk to-disk transfer as the rate limiting step for energy flow in phycobilisomes. Science **227**:419–423.
24. **Grossman, A. R., P. G. Lemaux, and P. B. Conley.** 1986. Regulated synthesis of phycobilisome components. Photochem. Photobiol. **44**:827–837.
25. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.

26. **Holschuk, K., W. Bottomley, and P. R. Whitfeld.** 1984. Structure of the spinach chloroplast genes for D2 and 44kd reaction-centre proteins of photosystem II and for $tRNA^{ser}$ (UGA). Nucleic Acids Res. **12**:8819–8834.
27. **Houmard, J., D. Mazel, C. Moguet, D. A. Bryant, and N. Tandeau de Marsac.** 1986. Organization and nucleotide sequence of genes encoding core components of the phycobilisomes from *Synechococcus* 6301. Mol. Gen. Genet. **205**:404–410.
28. **Hughes, E. O., P. R. Gorham, and A. Zehnder.** 1958. Toxicity of a unialgal culture of *Microcystis aeruginosa*. Can. J. Microbiol. **4**:225–236.
29. **Hunkapillar, W. M., E. Lujon, F. Ostrander, and L. E. Hood.** 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. **91**:227–236.
30. **Lemaux, P. G., and A. R. Grossman.** 1985. Major light-harvesting polypeptides encoded in polycistronic transcript in eukaryotic algae. EMBO J. **4**:1911–1919.
31. **Lemaux, P. G., and A. R. Grossman.** 1984. Isolation and characterization of a gene for a major light-harvesting polypeptide from *Cyanophora paradoxa*. Proc. Natl. Acad. Sci. USA **81**:4100–4104.
32. **Lundell, D. J, and A. N. Glazer.** 1983. Molecular architecture of light harvesting antenna. Quaternary interactions in the *Synechococcus* 6301 phycobilisome core as revealed by partial tryptic digestion and circular dichroism studies. J. Biol. Chem. **258**:8708–8713.
33. **Lundell, D. J., R. C. Williams, and A. N. Glazer.** 1981. Molecular architecture of light harvesting antenna. In vitro assembly of the rod substructures of *Synechococcus* 6301 phycobilisomes. J. Biol. Chem. **256**:3580–3592.
34. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Mazel, D., G. Guglielmi, J. Houmard, W. Sidler, D. A. Bryant, and N. Tandeau de Marsac.** 1986. Green light induces transcription of phycoerythrin operon in the cyanobacterium *Calothrix* 7601. Nucleic Acids Res. **14**:8279–8290.
36. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **110**:20–78.
37. **Nicholson, B. J., M. W. Hunkapiller, L. B. Grim, L. E. Hood, and J.-P. Revel.** 1981. Rat liver gap junction protein: properties and partial sequence. Proc. Natl. Acad. Sci. USA **78**:7594–7598.
38. **Redlinger, T., and E. Gantt.** 1982. An M_r 95,000 polypeptide in *Porphyridium cruentum* phycobilisomes. Possible function in linkage of phycobilisomes to thylakoids and in energy transfer. Proc. Natl. Acad. Sci. USA **79**:5542–5546.
39. **Redlinger, T., and E. Gantt.** 1981. Phycobilisome structure of *Porphyridium cruentum*. Polypeptide composition. Plant Physiol. **68**:1375–1399.
40. **Saraste, M.** 1984. Location of haem-binding sites in mitochondrial cytochrome *c*. FEBS Lett. **166**:367–372.
41. **Schirmer, T., W. Bode, R. Huber, W. Sidler, and H. Zuber.** 1985. X-ray crystallographic structure of the light harvesting biliprotein c-phycocyanin from the thermophilic cyanobacterium *Mastigocladus laminosus* and its resemblance to globin structure. J. Mol. Biol. **184**:257–277.
42. **Singer, B. S., L. Gold, S. T. Shinedling, L. R. Hunt, D. Pribnow, and M. A. Nelson.** 1981. Analysis in vivo of translational mutants of the rIIB cistron of bacteriophage T4. J. Mol. Biol. **149**:405–432.
43. **Sproat, B. S., and M. J. Gait.** 1984. Solid phase synthesis of oligodeoxyribonucleotides by the phosphotriester method, p. 83–115. *In* M. J. Gait (ed.), Oligonucleotide Synthesis. IRL Press Ltd., Oxford.
44. **Stormo, G. D.** 1986. Translation initiation, p. 195–224. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, London.
45. **Tandeau de Marsac, N., and G. Cohen-Bazire.** 1977. Molecular composition of cyanobacterial phycobilisomes. Proc. Natl. Acad. Sci. USA **74**:1635–1639.
46. **Thomas, P.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
47. **Williams, R. C., J. C. Gingrich, and A. N. Glazer.** 1980. Cyanobacterial phycobilisomes. Particles of *Synechocystis* 6701 and two pigment mutants. J. Cell Biol. **85**:558–566.
48. **Wu, R., N.-H. Wu, Z. Hanna, F. Georges, and S. Narang.** 1984. Purification and sequence analysis of synthetic oligodeoxyribonucleotides, p. 135–151. *In* J. M. Gait (ed.), Oligonucleotide Synthesis. IRL Press Ltd., Oxford.
49. **Zuber, H.** 1986. Structure of light harvesting antenna complexes of photosynthetic bacteria, cyanobacteria and red algae. Trends Biochem. Sci. **11**:414–419.