Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 6983–6987, November 1984
Biochemistry

Cloning and sequencing of the genes encoding the α and β subunits of C-phycocyanin from the cyanobacterium *Agmenellum quadruplicatum*

(oligonucleotide probe/DNA sequence/protein sequence homologies/blue-green algae/phycobiliprotein)

TAMI J. PILOT* AND J. LAWRENCE FOX†

Department of Zoology, The University of Texas, Austin, TX 78712

Communicated by E. E. Snell, July 26, 1984

ABSTRACT **Synthetic oligonucleotide probes were used to identify a cloned DNA fragment from the cyanobacterium *Agmenellum quadruplicatum* that contains the genes for the α and β subunits of C-phycocyanin. The coding region for the α-subunit gene begins 108 base pairs downstream from the 3′ end of the β-subunit structural gene. The sequences of the coding regions for both genes have been determined as well as 379 base pairs of 5′ flanking region, 204 base pairs of 3′ flanking region, and the 108 base pairs between the two genes. The site of transcriptional initiation is located ≈325 base pairs upstream from the β-subunit gene, and an open reading frame 114 base pairs long is found within this region. The significance of this additional open reading frame is not yet known. The derived amino acid sequences for both C-phycocyanin subunits were compared with other known C-phycocyanin sequences for homology. Homologies between the *A. quadruplicatum* α subunit and α subunits from other species were ≈70%, as were homologies between the *A. quadruplicatum* β subunit and other β subunits. Homologies between the various α and β subunits were 21%–27%. Codon usage for both the C-phycocyanin α- and β-subunit genes shows asymmetries for many amino acids that correspond closely to those seen in highly expressed *Escherichia coli* genes.**

Phycobiliproteins are protein–chromophore conjugates found in cyanobacteria (blue-green algae) as well as in rhodophytes (red algae) and cryptomonads. They are accessory photosynthetic pigments and may comprise as much as 40%–60% of the soluble protein in these cells (1–3). Thus, phycobiliproteins are major metabolic products of these organisms. These proteins can be divided into three classes based on their spectral properties: phycoerythrins, phycocyanins, and allophycocyanins. The basic structural unit of the proteins contains one α chain and one β chain. The α polypeptide is covalently linked to one bilin derivative; the β polypeptide is covalently linked to two bilin derivatives. *In vivo*, phycobiliproteins exist in higher aggregation states with the $(\alpha\beta)_6$ form generally believed to be the functionally important aggregate (3–6).

The role of phycobiliproteins is to trap light energy in the 500- to 650-nm wavelength range and to transfer it to chlorophyll a of photosystem II. *In vivo*, this transfer of energy occurs with an efficiency approaching 100% (7). Measurements of energy transfer in both intact algal cells and isolated phycobilisomes (8–11) have shown that light energy is normally first absorbed by phycoerythrin, then transferred to phycocyanin, then to allophycocyanin, and finally to chlorophyll a. It has further been established that the transfer of energy within a phycobiliprotein is from the β chain to the α chain (12–14).

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.

We are interested in the structure and function of phycobiliproteins from the cyanobacterium *Agmenellum quadruplicatum* (strain PR6) as well as in the organization and regulation of the genes that encode these proteins. This organism lacks phycoerythrin entirely and produces only small amounts of allophycocyanin. Thus, C-phycocyanin (C-PC) is the predominant phycobiliprotein in these cells. We report here the nucleotide sequences of the genes for both subunits of *A. quadruplicatum* C-PC as well as the sequence of a 114-base-pair (bp) open reading frame located 112 bp upstream from the 5′ end of the β-subunit gene. We also report the derived amino acid sequences and compare these with available amino acid sequences for C-PC from several other algal species.

Until now, no nucleotide sequence data existed for any phycobiliprotein genes. Sequence analysis at the nucleotide level allows us to determine the structural relationship of the individual subunit genes within the genome. It also allows us to search the regulatory sequences of these genes for any unique features. As more nucleotide sequences for phycobiliprotein genes from other organisms become available, homology studies can be undertaken at the nucleotide level.

MATERIALS AND METHODS

Library Construction. Genomic DNA from *A. quadruplicatum* was partially digested with *Sau*3A and size-fractionated on a Sepharose 4B column. Fragments ranging from 2 to 10 kilobase pairs (kbp) were ligated into pUC9 (15), which had been digested with *Bam*HI. This DNA was used to transform *Escherichia coli* JM83 (16). Approximately 8000 recombinant colonies were isolated for screening.

Probes. Two oligonucleotides were synthesized by the phosphoramidite method (17–19) for use as hybridization probes. The sequence of the first probe, 5′-A-T-G-T-T-$\substack{C\\T}$-G-A-$\substack{C\\T}$-A-T-$\substack{C\\T\\A}$-T-T, was based on sequence data for the first five amino acids of the C-PC β subunit (20). The sequence of the second probe, 5′-G-A-$\substack{C\\T}$-A-T-G-G-A-$\substack{A\\G}$-A-T-$\substack{C\\T\\A}$-A-T, was based on sequence data for amino acids 87–91 of the β subunit (unpublished data).

Screening. Recombinant colonies were transferred to nitrocellulose filters (21) and then hybridized by a modifica-

Abbreviations: kbp, kilobase pair(s); bp, base pair(s); C-PC, C-phycocyanin.

*Present address: Department of Molecular Biology, Abbott Laboratories, North Chicago, IL 60064.
†Present address: Department of Molecular Biology, Abbott Laboratories, North Chicago, IL 60064; and Department of Biochemistry, Molecular and Cell Biology, Northwestern University, Evanston, IL 60201.

tion of the Grunstein and Hogness method (22): nitrocellulose filters with bound colony DNA were prehybridized at 55°C for 2–3 hr in a large beaker containing 10× Denhardt's medium (0.2% bovine serum albumin/0.2% Ficoll/0.2% polyvinylpyrrolidone) and 6× NaCl/Cit buffer (0.9 M NaCl/0.09 M sodium citrate, pH 7.0) at 5 ml per filter. Filters were then transferred to Seal-And-Save bags (Sears) containing 10× Denhardt's medium, 6× NaCl/Cit at 3 ml per filter, and $\approx 10^6$ cpm of labeled probe per filter. Probes were 5' end-labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase. Filters were hybridized overnight at 37°C, then removed from the bags and washed 3 times in 2× NaCl/Cit for 10 min at room temperature. After washing, filters were allowed to dry and were then exposed to x-ray film overnight with an intensifying screen.

Restriction and Sequence Analysis. Plasmid DNA was subjected to restriction analysis and Southern blot hybridizations (23) with the two ^{32}P-labeled oligonucleotide probes. Hybridization conditions were essentially the same as those used for the colony screening. Sequencing was by the base-specific chemical cleavage method of Maxam and Gilbert (24).

RNA Isolation. Total *A. quadruplicatum* RNA was isolated by a modification of the method of Chirgwin *et al.* (25). One-liter *A. quadruplicatum* cultures were grown at 39°C in medium A (26) supplemented with 5% CO_2/95% air. Cells were harvested at late logarithmic phase, washed with 0.12 M NaCl/0.05 M EDTA, pH 8.0, then resuspended in 20 ml of RNA lysis buffer (4 M guanidinium isothiocyanate/0.5% sodium *N*-lauroylsarcosine/0.1 M 2-mercaptoethanol/1 mM EDTA/0.2 M sodium acetate, pH 5.2) (25), and then passed through a French pressure cell at 10,000 psi (1 psi = 6.895×10^3 Pa). Lysates were then layered into ultracentrifuge tubes one-quarter filled with 5.7 M cesium chloride/0.05 M sodium acetate, pH 5.2/1 mM EDTA, and centrifuged for 16–18 hr in a Beckman SW27 rotor at 25,000 rpm and 20°C. The RNA pellets were redissolved in RNA lysis buffer, ethanol-precipitated, resuspended in 0.3 M sodium acetate (pH 5.2), ethanol-precipitated again, and then resuspended in H_2O and stored at −70°C.

S1 Nuclease Mapping. The 5' end of the mRNA for the *A. quadruplicatum* C-PC genes was mapped by a modification of the procedure from Weaver and Weissmann (27). Single-stranded end-labeled fragments were prepared by labeling the 5' ends with T4 polynucleotide kinase and electrophoresis through strand separation gels (24). Approximately 50,000 cpm of the labeled fragments and 50 μg of total *A. quadruplicatum* RNA in 10 μl of 35% formamide/0.4 M NaCl/1 mM EDTA/0.04 M Pipes, pH 6.4, were sealed in glass capillaries, heated at 65°C for 10 min, and then incubated for 12–14 hr at 42°C. Samples were diluted into cold S1 nuclease buffer (27), digested with various amounts of S1 nuclease for 30 min, and run on denaturing polyacrylamide/urea gels (24).

RESULTS AND DISCUSSION

Location and Sequence of *A. quadruplicatum* C-PC Genes. One clone containing a 4.1-kbp insert was chosen for further analysis. Plasmid DNA (designated pTP1) from this clone was digested with *Hinc*II, and the fragments were hybridized with each of the oligonucleotide probes. A 0.6-kbp fragment that hybridized with probe 2, but not with probe 1, was isolated and the ends were sequenced to confirm that pTP1 did contain the C-PC genes. One end was found to lie within the C-PC β-subunit gene; the other end was located within the α-subunit gene. This conclusion was based on the putative DNA sequences derived from reverse transcription of the amino acid sequences. A detailed restriction map of the region was generated, and the region containing the C-PC genes was sequenced (Fig. 1).

Fig. 2 displays the noncoding strand sequences of the C-PC β-subunit and α-subunit genes. The β-subunit gene is upstream from the α-subunit gene, with a 108-bp segment sepa-

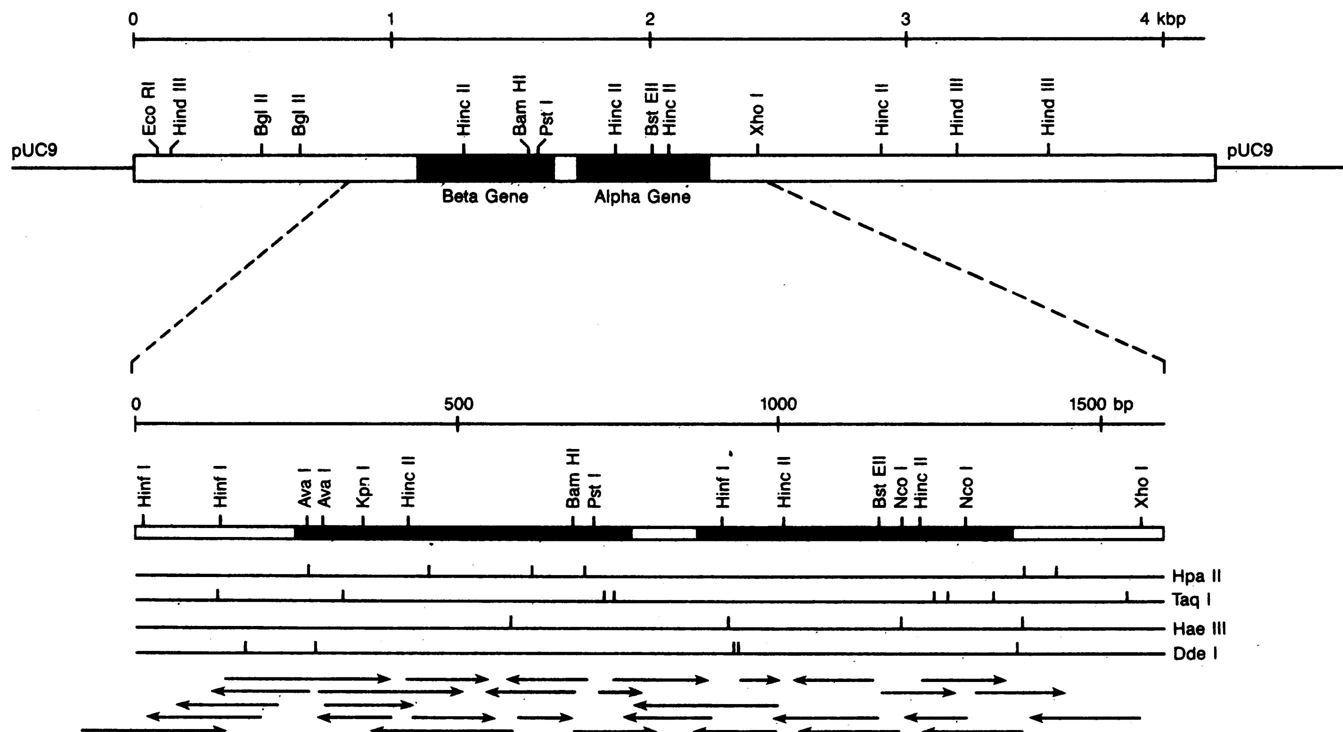

FIG. 1. Restriction map and sequence analysis strategy for the C-PC α- and β-subunit genes of *A. quadruplicatum*. A map of the cloned 4.1-kbp fragment (*Upper*) and an expanded map of the 1.6-kbp region (*Middle*) subjected to sequence analysis are shown. The arrows (*Lower*) indicate direction and extent of sequence analysis.

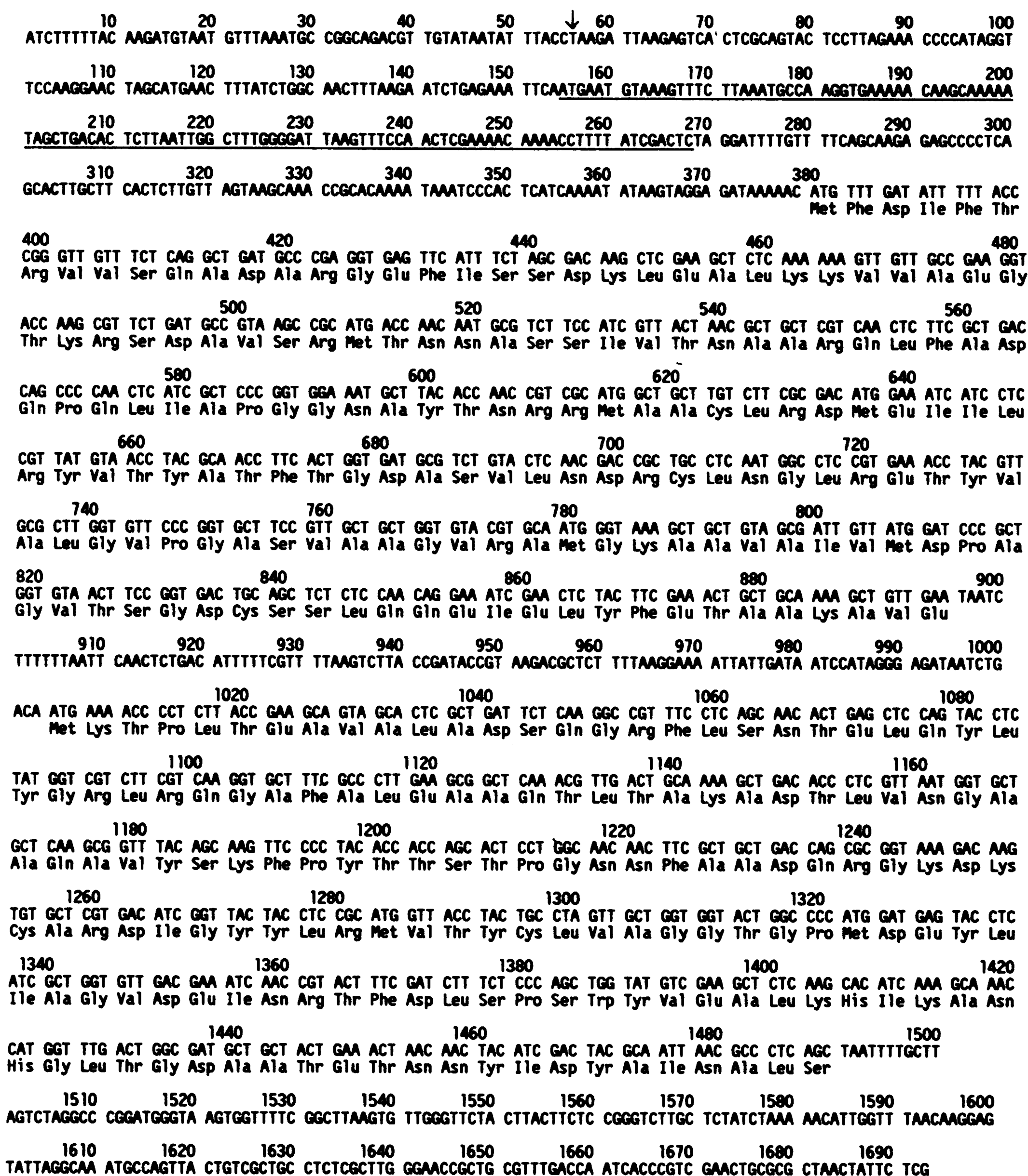

FIG. 2. Nucleotide sequences of *A. quadruplicatum* C-PC α-subunit and β-subunit genes and flanking regions. The β-subunit gene encompasses bases 380–895, and the α-subunit gene includes bases 1004–1489. The 114-bp open reading frame (bases 155–268) is underlined. Arrow indicates the identified site of transcription initiation (plus or minus $\approx$5 bases). The derived amino acid sequences of the α and β subunits are also shown.

rating them. The β-subunit gene-coding region contains 516 bp, and the α-subunit gene-coding region contains 486 bp.

The 5′ end of the C-PC transcript was located by S1 nuclease mapping (Fig. 3). The 0.4-kbp *Bgl* II/*Ava* I fragment containing 17 bp of β-subunit gene-coding region and 380 bp of 5′ flanking sequence was isolated, labeled, and strand separated. The coding strand was hybridized with total *A. quadruplicatum* RNA, treated with S1 nuclease, and run on a denaturing polyacrylamide/urea gel. The major protected fragment indicates that transcription initiates $\approx$325 bp upstream from the 5′ end of the β-subunit gene. A minor fragment $\approx$180 bp long is also seen after S1 nuclease treatment, but we believe this is an artifact resulting from melting of the duplex at this position in the sequence. This region (bases 213–224) is highly A+T-rich. Just upstream from the identified initiation site is a sequence identical to the consensus prokaryotic RNA polymerase binding site, T-A-T-A-A-T (Fig. 2; bases 43–49) (28). No other sequences resembling this consensus sequence are found anywhere upstream from the β-subunit or α-subunit genes.

In most prokaryotes, the 5′ untranslated region on mRNA molecules is generally not more than 200 bp long (29). Because we found that transcription initiates $\approx$325 bp upstream from the β-subunit gene, we examined this region for unusual features. We found an open reading frame starting with an ATG that is 114 bp long. It starts $\approx$100 bp downstream from the transcription initiation site and ends 112 bp upstream from the 5′ end of the β-subunit gene (underlined in Fig. 2).

The sequences just upstream from the β-subunit and α-subunit genes as well as the sequence upstream from the 114-bp open reading frame were examined for possible ribosome-binding sites. The sequence G-A-G-A is found 8–11 bp upstream from the β-subunit gene, 11–14 bp upstream from the α-subunit gene, and 7–10 bp upstream from the 114-bp open reading frame. This sequence is similar, although not identical, to the consensus prokaryotic ribosome binding sequence

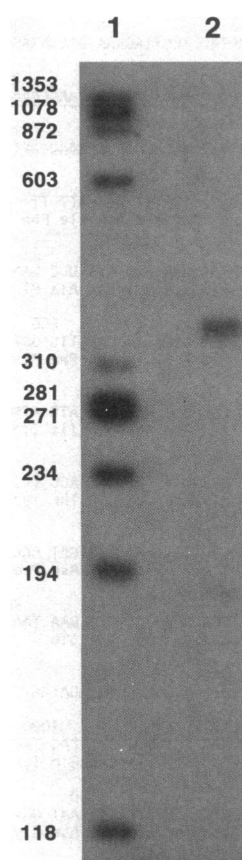

FIG. 3. S1 nuclease mapping of the 5′ end of the C-PC transcript. The 0.4-kbp *Bgl* II/*Ava* I fragment containing 17 bp of β-subunit gene-coding region and 380 bp of 5′ flanking sequence was isolated and then labeled and strand-separated. Coding strand was hybridized with total *A. quadruplicatum* RNA as described, then treated with 20 units of S1 nuclease, and run on a 5% denaturing polyacrylamide/urea gel (lane 2). Size standards in lane 1 are end-labeled *Hae* III-digested φX174 fragments.

G-G-A-G (30). The significance of this cannot be confirmed, however, because the sequence of *A. quadruplicatum* 16S rRNA has not been determined.

RNA blot hybridizations were carried out to confirm that the β-subunit and α-subunit genes are transcribed on the same mRNA (unpublished data). Probes specific for the β-subunit gene, the α-subunit gene, and the region upstream from the β-subunit gene all hybridized to a single mRNA of the same size, 1450–1500 bases long. This is the expected length for a message that initiates at the identified transcription initiation site ≈325 bp upstream from the β-subunit gene and continues through the end of the α-subunit gene. Thus, it appears that this transcript contains the C-PC subunit genes, but not the genes for the allophycocyanin subunits or any other genes.

It has not been determined whether the 114-bp open reading frame upstream from the β-subunit gene is translated into a polypeptide, but its distance from the initiation site and the fact that it has a possible ribosome binding site indicate that it could be. It is in the same reading frame as the β-subunit gene, but 5 termination codons lie in between, so it does not appear to be a leader sequence. This polypeptide, if translated, may be involved in phycocyanobilin synthesis or attachment, in formation of phycobilisomes, or in the regulation of the C-PC genes. Further genetic analysis of *A. quadruplicatum* is needed before this can be determined.

Amino Acid Sequences of *A. quadruplicatum* C-PC Subunits. The predicted amino acid sequences of the C-PC β and α subunits are shown in Fig. 2. The β subunit contains 172 amino acids and has a calculated molecular weight of 18,300; the α subunit contains 162 amino acids and has a calculated molecular weight of 17,600. The derived sequences agree closely with the amino acid sequence data generated by protein sequencing (unpublished data). These amino acid sequences were compared for homology with known C-PC sequences from three other algal species. The results are shown in Table 1.

Table 1. % sequence homologies between C-PC subunits of *A. quadruplicatum* and other organisms

	Agmenellum		*Anacystis*		*Mastigocladus*		*Cyanidium*	
	α	β	α	β	α	β	α	β
α	—	22.8	74.7	26.5	69.8	25.3	74.7	25.9
β	—	—	23.5	70.3	21.0	64.5	21.6	65.7

Data are from the following references: *Anacystis nidulans*, α subunit (31), β subunit (32); *Mastigocladus laminosus* (33); *Cyanidium caldarium*, α subunit (34), β subunit (35).

Overall homologies between *A. quadruplicatum* C-PC α

Table 2. Codon usage in *A. quadruplicatum* C-PC α-subunit and β-subunit genes

		α	β			α	β			α	β			α	β
UUU	Phe	0	2	UCU	Ser	2	6	UAU	Tyr	2	1	UGU	Cys	1	1
UUC	Phe	5	4	UCC	Ser	0	3	UAC	Tyr	9	4	UGC	Cys	1	2
UUA	Leu	0	0	UCA	Ser	0	0	UAA	Ter	1	1	UGA	Ter	0	0
UUG	Leu	2	0	UCG	Ser	0	0	UAG	Ter	0	0	UGG	Trp	1	0
CUU	Leu	4	2	CCU	Pro	2	0	CAU	His	1	0	CGU	Arg	5	6
CUC	Leu	9	10	CCC	Pro	3	4	CAC	His	1	0	CGC	Arg	2	4
CUA	Leu	1	0	CCA	Pro	0	0	CAA	Gln	4	3	CGA	Arg	0	1
CUG	Leu	0	0	CCG	Pro	0	0	CAG	Gln	2	3	CGG	Arg	0	1
AUU	Ile	1	3	ACU	Thr	8	4	AAU	Asn	1	3	AGU	Ser	0	0
AUC	Ile	5	5	ACC	Thr	6	7	AAC	Asn	8	4	AGC	Ser	5	3
AUA	Ile	0	0	ACA	Thr	0	0	AAA	Lys	4	4	AGA	Arg	0	0
AUG	Met	3	6	ACG	Thr	1	0	AAG	Lys	3	2	AGG	Arg	0	0
GUU	Val	5	10	GCU	Ala	14	17	GAU	Asp	4	5	GGU	Gly	9	10
GUC	Val	1	0	GCC	Ala	2	3	GAC	Asp	6	5	GGC	Gly	4	1
GUA	Val	1	6	GCA	Ala	5	3	GAA	Glu	5	8	GGA	Gly	0	1
GUG	Val	0	0	GCG	Ala	2	4	GAG	Glu	2	1	GGG	Gly	0	0

Ter, termination.

and the other known α sequences are 69%–75%; the homologies between *A. quadruplicatum* C-PC β and the other known β sequences are 64%–71%. The regions where the phycocyanobilin groups attach (α amino acid 84 and β amino acids 82 and 153) are highly conserved in both *A. quadruplicatum* C-PC subunits just as they are in the other species (36, 37).

It is believed that all phycobiliprotein genes arose from a single ancestral gene by gene duplication, and several models have been proposed to explain this (35, 38, 39). In these models, the genes for the allophycocyanin subunits are proposed to have arisen first, giving rise then to the genes for the C-PC subunits. Both subunits of allophycocyanin have only one attached chromophore, as does the α subunit of C-PC; the C-PC β subunit has two attached chromophores. In the model of Troxler *et al.* (35), this second site for chromophore attachment in the C-PC β subunit is due to the insertion of 12 amino acids at positions 146–157.

Thus, to properly align α and β subunits for homology comparisons, this second chromophore region must be looped out, along with amino acids 80–81 of the α chain. When this alignment is made, the homologies between C-PC α and the β subunits is 22%–27% and between C-PC β and the α subunits is 21%–23.5%. There are small regions of higher homology scattered throughout the entire lengths of the sequences, especially around the attachment site for the first chromophore, but there are no long homologous stretches.

Codon Usage in *A. quadruplicatum* C-PC Genes. Codon usage in both the β-subunit and α-subunit genes of *A. quadruplicatum* C-PC (Table 2) shows asymmetries that correspond closely to those seen in highly expressed *E. coli* genes (40). Although the abundance of various *A. quadruplicatum* tRNA species is not known, it would appear to be similar to that of *E. coli*, except for $tRNA^{Pro}$, because CCC is the preferred codon in *A. quadruplicatum* rather than CCG as in *E. coli*.

Another criterion used to judge which codon will be preferred for a particular amino acid in highly expressed genes is a bias in choice between cytosine and uracil in the third position. The bias is toward the nucleotide that yields a codon-anticodon binding energy of intermediate strength. Gouy and Gautier (40) show that if the first two bases are both adenine or uracil, then cytosine is preferred over uracil in position three, while if the first two bases are both cytosine or guanine, then uracil is preferred. The codons for the *A. quadruplicatum* C-PC genes adhere to this, except for those which code for proline. The codon CCC is seen 7 times in these two genes, while CCU only appears twice.

We thank Dr. E. Eileen Gardner and Mrs. Sandra Smith for their protein sequence determinations. We also thank Dr. Wlodek Mandecki, Dr. Jim Kramer, and Mr. Tim Bolling for technical assistance and helpful advice. This work was supported by Welch Foundation Grant F-630, National Institutes of Health Grant GM 21688, and Abbott Laboratories.

1. Myers, J. & Kratz, W. A. (1955) *J. Gen. Physiol.* **39**, 11–22.
2. Bennett, A. & Bogorad, L. (1971) *Biochemistry* **10**, 3625–3634.
3. Bogorad, L. (1975) *Annu. Rev. Plant Physiol.* **26**, 369–401.
4. Bryant, D. A., Guglielmi, G., Tandeau de Marsac, N., Castets, A. & Cohen-Bazire, G. (1979) *Arch. Microbiol.* **123**, 113–127.
5. Lundell, D. J., Williams, R. C. & Glazer, A. N. (1981) *J. Biol. Chem.* **256**, 3580–3592.
6. Gantt, E. (1980) *Int. Rev. Cytol.* **66**, 48–80.
7. Halldal, P. (1970) in *Photobiology of Microorganisms*, ed. Halldal, P. (Wiley, London), pp. 17–55.
8. Ley, A. C. & Butler, W. L. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3957–3960.
9. Grabowski, J. & Gantt, E. (1978) *Photochem. Photobiol.* **28**, 39–45.
10. Grabowski, J. & Gantt, E. (1978) *Photochem. Photobiol.* **28**, 47–54.
11. Porter, G., Tredwell, C. J., Searle, G. F. W. & Barber, J. (1978) *Biochim. Biophys. Acta* **501**, 232–245.
12. Glazer, A. N., Fang, S. & Brown, D. M. (1973) *J. Biol. Chem.* **248**, 5679–5685.
13. Csatorday, K. (1978) *Biochim. Biophys. Acta* **504**, 341–343.
14. MacColl, R. & Berns, D. S. (1978) *Photochem. Photobiol.* **27**, 343–349.
15. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
16. Mandel, M. & Higa, A. (1970) *J. Mol. Biol.* **53**, 159–162.
17. Matteucci, M. D. & Caruthers, M. H. (1980) *Tetrahedron Lett.* **21**, 719–722.
18. Matteucci, M. D. & Caruthers, M. H. (1981) *J. Am. Chem. Soc.* **103**, 3185–3191.
19. McBride, L. J. & Caruthers, M. H. (1983) *Tetrahedron Lett.* **24**, 245–248.
20. Gardner, E. E., Stevens, S. E. & Fox, J. L. (1980) *Biochim. Biophys. Acta* **624**, 187–195.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 318–319.
22. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
23. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
24. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
25. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
26. Stevens, S. E., Patterson, C. O. P. & Myers, J. (1973) *J. Phycol.* **9**, 427–430.
27. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **7**, 1175–1193.
28. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13**, 319–353.
29. Kozak, M. (1983) *Microbiol. Rev.* **47**, 1–45.
30. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981) *Annu. Rev. Microbiol.* **35**, 365–403.
31. Walsh, R. G., Wingfield, P., Glazer, A. N. & DeLange, R. J. (1980) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **39**, 1998 (abstr.).
32. Freidenreich, P., Apell, G. S. & Glazer, A. N. (1978) *J. Biol. Chem.* **253**, 212–219.
33. Frank, G., Sidler, W., Widmer, H. & Zuber, H. (1978) *Hoppe-Seyler's Z. Physiol. Chem.* **359**, 1491–1507.
34. Offner, G. D., Brown-Mason, A. S., Ehrhardt, M. M. & Troxler, R. F. (1981) *J. Biol. Chem.* **256**, 12167–12175.
35. Troxler, R. F., Ehrhardt, M. M., Brown-Mason, A. S. & Offner, G. D. (1981) *J. Biol. Chem.* **256**, 12176–12184.
36. Williams, V. P. & Glazer, A. N. (1978) *J. Biol. Chem.* **253**, 202–211.
37. Bryant, D. A., Hixson, C. S. & Glazer, A. N. (1978) *J. Biol. Chem.* **253**, 220–225.
38. Glazer, A. N. (1976) *Photochem. Photobiol. Rev.* **1**, 71–115.
39. Zuber, H. (1978) *Ber. Dtsch. Bot. Ges. Bd.* **91**, 459–475.
40. Gouy, M. & Gautier, C. (1982) *Nucleic Acids Res.* **10**, 7055–7074.