Photosynthesis Research 18:99–132(1988)

Minireview

Photoregulation of gene expression in the filamentous cyanobacterium *Calothrix* sp. PCC 7601: light-harvesting complexes and cell differentiation

NICOLE TANDEAU DE MARSAC, DIDIER MAZEL, THIERRY DAMERVAL, GÉRARD GUGLIELMI, VÉRONIQUE CAPUANO & JEAN HOUMARD

Unité de Physiologie Microbienne (C.N.R.S., U.A. 1129), Département de Biochimie et Génétique Moléculaire, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

Received 6 October 1987; Accepted 23 December 1987

Key words: *Calothrix* sp. PCC 7601, gas vesicles, hormogonia, photoregulation, phycobiliproteins, phycobilisomes

Abstract. Light plays a major role in many physiological processes in cyanobacteria. In *Calothrix* sp. PCC 7601, these include the biosynthesis of the components of the light-harvesting antenna (phycobilisomes) and the differentiation of the vegetative trichomes into hormogonia (short chains of smaller cells). In order to study the molecular basis for the photoregulation of gene expression, physiological studies have been coupled with the characterization of genes involved either in the formation of phycobilisomes or in the synthesis of gas vesicles, which are only present at the hormogonial stage.

In each system, a number of genes have been isolated and sequenced. This demonstrated the existence of multigene families, as well as of gene products which have not yet been identified biochemically. Further studies have also established the occurrence of both transcriptional and post-transcriptional regulation. The transcription of genes encoding components of the phycobilisome rods is light-wavelength dependent, while translation of the phycocyanin genes may require the synthesis of another gene product irrespective of the light regime. In this report, we propose two hypothetical models which might be part of the complex regulatory mechanisms involved in the formation of functional phycobilisomes. On the other hand, transcription of genes involved in the gas vesicles formation (*gvp* genes) is turned on during hormogonia differentiation, while that of phycobiliprotein genes is simultaneously turned off. In addition, an antisense RNA which might modulate the translation of the *gvp* mRNAs is synthezised.

Abbreviations: AP – allophycocyanin, APB – allophycocyanin B, bp – base pair, GVP – gas vesicle protein, kb – kilobase, kDa – kilodalton, L – linker polypeptide, PC – phycocyanin, PCC – Pasteur Culture Collection, PE – phycoerythrin, UTEX – University of Texas algal collection

1. Introduction

In cyanobacteria, like in higher plants, light is not only used to perform photosynthesis but also to trigger and modulate various developmental and regulatory processes (Stanier and Cohen-Bazire 1977). The best known examples of photoregulation in cyanobacteria are those which affect pigment content and cell differentiation. No photoreceptor involved in such processes has yet been isolated in these prokaryotic organisms, although physiological evidence strongly suggests that such molecules exist to mediate responses to light (for reviews, see Tandeau de Marsac 1983 and Grossman et al. 1986). The purpose of this review is to present our recent progress in this research area. No attempt has been made to extensively cover the literature but we have chosen to discuss the molecular aspects since most of the physiological and biochemical studies performed in this field have often been reviewed during the last decade. When appropriate, the reader will be directed to specialized recent reviews in which more detailed information and original references are available.

In contrast to higher plants, cyanobacteria contain chlorophyll *a* but not chlorophyll *b*, and, like in rhodophyta, phycobiliproteins constitute the major light-harvesting antenna. These chromoproteins are organized into multimolecular structures called phycobillisomes which, although functionally analogous to the light-harvesting chlorophyll-protein complexes of higher plants, are attached to instead of being embedded in the photosynthetic membranes. Phycobilisomes are composed of two domains: the central core proximal to the photosynthetic membrane, and six rods which radiate from the core. This generally results in a hemidiscoidal structure. In addition, linker polypeptides, which are specifically associated with each phycobiliprotein complex of the phycobilisome, maintain its physical integrity and slightly modify the intrinsic spectral characteristics of the constituent phycobiliproteins, resulting in an optimization of the energy transfer to the terminal acceptor(s) of the phycobilisome. This structure is thus particularly well adapted to the trapping and funnelling of the photon energy to the photosystem II reaction centers (for reviews on phycobiliproteins and phycobilisomes, see for example, Glazer 1984 and 1985).

It is known that in general, in photosynthetic organisms, antenna size is inversely proportional to the light intensity received by the cultures during growth. Similarly, cyanobacteria respond to light intensity by increasing (under low light intensity) or decreasing (under high light intensity) their chlorophyll *a* and phycobiliprotein content, but the mechanisms by which light regulates these processes remain unknown. In contrast, the phenomenon of complementary chromatic adaptation, in which the syn-

thesis of some components of the phycobilisome rods can be regulated by the incident light wavelength, has been more extensively studied and is undoubtedly governed by at least one photoreversible pigment. This photoreceptor presents some analogy with the phytochrome of higher plants, although the most efficient wavelengths for its interconversion are different (green versus red radiation, instead of red versus far-red radiation for phytochrome; for a review on complementary chromatic adaptation, see Tandeau de Marsac 1983).

In some filamentous cyanobacterial strains, light also controls differentiation processes such as hormogonia differentiation. This process, which is part of a developmental cycle, is a prerequisite for the dispersal of these species in their natural habitats. Indeed, in contrast to vegetative trichomes from which they differentiate, hormogonia are more resistant to environmental stresses; they are also motile, the motility being correlated with the appearance of pili at their cell surface (G. Guglielmi unpublished data) and they are filled with gas vesicles which provide them with buoyancy (Walsby 1981). Interestingly, this differentiation process has been shown to depend upon several environmental factors, among which are red radiations (Armstrong et al. 1983, Tandeau de Marsac 1983, Rippka and Herdman 1985). Whether such a phenomenon is controlled by the same photoreceptor as the one involved in complementary chromatic adaptation is an open question.

With the aim of understanding the molecular basis of these physiological adaptations to the environment, and to determine if one or several photoreceptors are involved in these regulatory processes, we have undertaken the characterization of the structure, organization, and expression of the genes encoding phycobilisome components, as well as those involved in hormogonia differentiation in *Calothrix* 7601 (also called *Fremyella diplosiphon* UTEX 481)*. This filamentous cyanobacterium belongs to chromatic adapters of group III (Tandeau de Marsac 1983).

2. The phycobilisome components: phycobiliproteins and linker polypeptides

2.1 Features emerging from biochemical and physiological data

Biochemical and physiological studies have shown that phycobiliproteins are stable chromoproteins, which can represent up to 50% of the total cell protein content and are organized in phycobilisomes. In response to changes

*Cyanobacterial strains are designated throughout this review by their genus name followed by their number in the Pasteur Culture Collection when available (for example, *Calothrix* 7601 = *Calothrix* sp. PCC 7601).

of the incident light, the number of phycobilisomes per cell, as well as the molecular composition and the size of the rods of the phycobilisomes, may vary (Raps et al. 1985; for reviews, see Cohen-Bazire and Bryant 1982, Glazer 1982, Tandeau de Marsac 1983).

Analysis of *Calothrix* 7601 phycobilisomes, purified from fully red- or green-light-adapted cells, has allowed a rather precise determination of their structure and composition under these two different light regimes. The core of both red- and green-light phycobilisomes seems invariant and is composed of at least five chromophoric polypeptides and one non-chromophoric linker polypeptide: the allophycocyanin subunits (αAP and βAP), allophycocyanin B (αAPB), the 18.3 kDa polypeptide (β18.3), the 92 kDa anchor polypeptide (L_{CM}^{92}) and the AP-associated linker ($L_C^{7.8}$). By analogy with *Synechocystis* 6701 phycobilisomes (Gingrich et al. 1983, Glazer and Clark 1986), it is thought that these polypeptides are organized into 12 trimers within the core of *Calothrix* 7601, the total number of molecules being 32, 34, 2, 2, 2, 6 for αAP, βAP, αAPB, β18.3 L_{CM}^{92} and $L_C^{7.8}$, respectively. In contrast to the core, the rods of red- and green-light phycobilisomes of *Calothrix* 7601 differ in composition (Tandeau de Marsac 1983, G. Guglielmi unpublished data). Rods of red-light phycobilisomes are composed of at least four chromophoric polypeptides: the phycocyanin-1 subunits (αPC1 and β PC1) and the phycocyanin-2 subunits (αPC2 andβPC2), the absorption maxima of which are situated at approximately 620 nm. These polypeptides are organized into hexamers $(\alpha\beta)6$. In addition, there are three PC-associated rod linker polypeptides (L_R^{38}, L_R^{39} and $L_R^{9.7}$) and one PC-associated core-rod linker polypeptide (L_{CR}^{30}). In rods of green-light phycobilisomes, the basal disc is a PC hexamer-L_{CR}^{30} complex like in red-light phycobilisome rods, while the more distal discs are PE hexamers whose absorption maxima are situated at approximately 565 nm. Two specific rod linker polypeptides (L_R^{35} and L_R^{36}) are associated with these hexamers of PE. A schematic representation of red- and green-light phycobilisomes from *Calothrix* 7601 is presented in Fig. 1.

No precise analysis has been performed on the composition of phycobilisomes from *Calothrix* 7601 cells grown under different light intensities. However, from information obtained with other cyanobacteria (Yamanaka and Glazer 1981, Raps et al. 1985), we anticipate that the number of phycobilisomes will be greater, and/or the size of their rods longer, in cells grown under low light intensity than in cells grown under high light intensity. In addition, a change in the PC/PE ratio could occur, if the level of synthesis of these two phycobiliproteins is not modified to the same extent in response to a given light intensity. A specific degradation of phycobilisome components, mainly PC, has been shown to occur in several *Anabaena* species (Foulds and Carr 1977, Wood and Haselkorn 1980), in

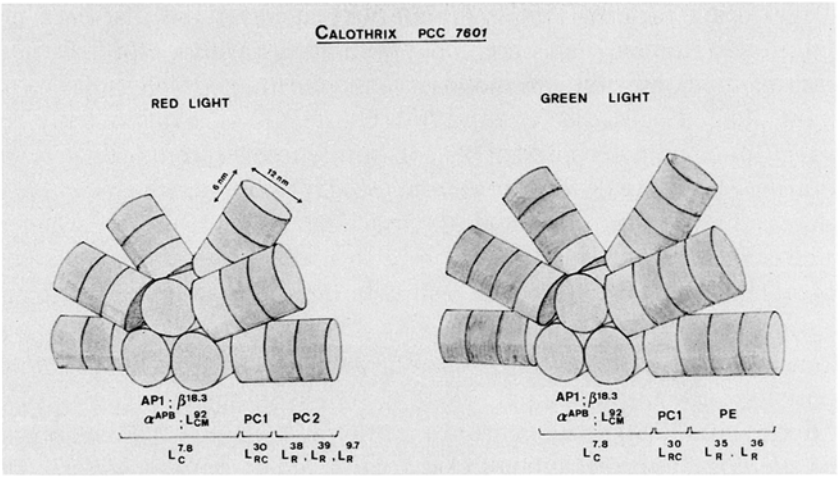

Fig. 1. Schematic representation of the *Calothrix* 7601 phycobilisomes purified from cells grown under red- or green-light conditions. For phycobiliproteins, the abbreviations AP1, PC1, PC2 and PE refer to the α and β subunits of allophycocyanin-1, phycocyanin-1, phycocyanin-2 and phycoerythrin, respectively. αAPB and β18.3 denote the α-type allophycocyanin B and the β-type phycobiliprotein of MM 18.3 kDa, respectively. Linker polypeptides are abbreviated, with a superscript denoting their apparent molecular mass in kDa and a subscript that specifies their location in the phycobilisome: R, rod substructure; RC, rod-core junction; C, core. The anchor polypeptide (MM 92 kDa) is denoted L^{92}_{CM}, with CM for core-membrane junction. The size of a disk (6 × 12 nm) corresponds to an hexamer $(\alpha\beta)$ 6 of PC or PE in the rod. The location of the linker polypeptides within the rods has not yet been precisely established.

Synechococcus 6301 (Lau et al. 1977, Yamanaka and Glazer 1980), *Synechococcus* 7002 (Stevens et al. 1981) and *Synechocystis* 6803 (Elmorjani and Herdman 1987) in response to nitrogen starvation. In contrast, adaptation of cyanobacterial cells to changes in light wavelength, and most probably in light intensity, does not result in the induction of specific degradative processes (Bennett and Bogorad 1973). This suggests that most of the events which lead to changes in the phycobilisome content of the cells, or in their composition, are very likely to result from precise transcriptional (and/or translational) controls of the expression of the genes encoding phycobilisome components.

2.2 Functional organization of the genes involved in the formation of phycobilisomes

2.2.1 Gene isolation and characterization

Comparisons of partial or complete amino acid sequences of phycobiliproteins have shown that individual subunits are highly conserved among

different cyanobacterial species (about 80% identity) and that each phycobiliprotein subunit shares homology with all the others, especially in the region of the conserved chromophore attachment site (Cohen-Bazire and Bryant 1982, Füglistaller et al. 1983, Glazer 1984). Consequently, our strategy for cloning the different phycobiliprotein genes from *Calothrix* 7601 was mostly based on the use of heterologous DNA probes which include the region coding for the chromophore attachment site. However, when the amino acid sequence was known, we also used synthetic oligonucleotide probes. Hybridization analyses of restriction enzyme digests were performed by using an *apcAB* probe (αAP and βAP) from *Synechococcus* 6301 (Houmard et al. 1986), *cpc*A (αPC) and *cpc*B (βPC) probes from *Synechococcus* 7002 (De Lorimier et al. 1984), as well as synthetic oligonucleotides (17 nucleotides long) designed from a portion of the amino acid sequence of the *Calothrix* 7601 αPE subunit (Sidler et al. 1986). Seven *Calothrix* 7601 *Eco*RI fragments of 9.5, 7.0, 6.5, 6.3, 4.5, 3.7 and 3.5 kb, cloned either from a λ EMBL3 genomic library or from partial libraries constructed in *E. coli* plasmid vectors, were totally or partially sequenced (for references, see below). The predicted amino acid sequence of each open reading frame has been compared with the sequences previously determined for proteins purified from *Calothrix* 7601 or from other cyanobacteria. Their transcription has been examined by hybridization experiments of total RNA extracted from *Calothrix* 7601 cells grown under green- or red-light conditions. Most of these open reading frames have thus been identified. As shown in Fig. 2, each of the cloned *Eco*RI fragments carries one or two genes encoding phycobiliprotein subunits and, for three of them, genes encoding linker polypeptides have been found. Some additional open reading frames, which do not correspond to any known phycobilisome components, were found to be involved in the formation of functional phycobilisome rods.

Phycobiliprotein genes. As expected from hybridization experiments, the 7.0 kb *Eco*RI fragment carries two genes, *cpeB* (552 bp) and *cpeA* (492 bp), which encode βPE and αPE, respectively (Mazel et al. 1986). These two genes are 79 nucleotides apart. It is surprising that, although only two different PCs (the 'constitutive' PC1 and 'inducible' PC2) have been characterized biochemically, three *Eco*RI fragments were found to harbor complete sets of PC genes. The first PC gene cluster, *cpcB1* (516 bp) and *cpcA1* (486 bp), is carried by the 4.5 kb *Eco*RI fragment and encodes βPC1 and αPC1, respectively (Conley et al. 1986, Mazel et al. 1988). The second PC gene cluster, *cpcB2* (516 bp) and *cpcA2* (486 bp), carried by the 6.5 kb *Eco*RI fragment, encodes βPC2 and αPC2, respectively (Conley et al. 1985, V. Capuano et al. 1988). The third PC gene cluster, *cpcB3* (519 bp) and *cpcA3* (486 bp) carried by the 9.5 kb *Eco*RI fragment encodes the β and α subunits

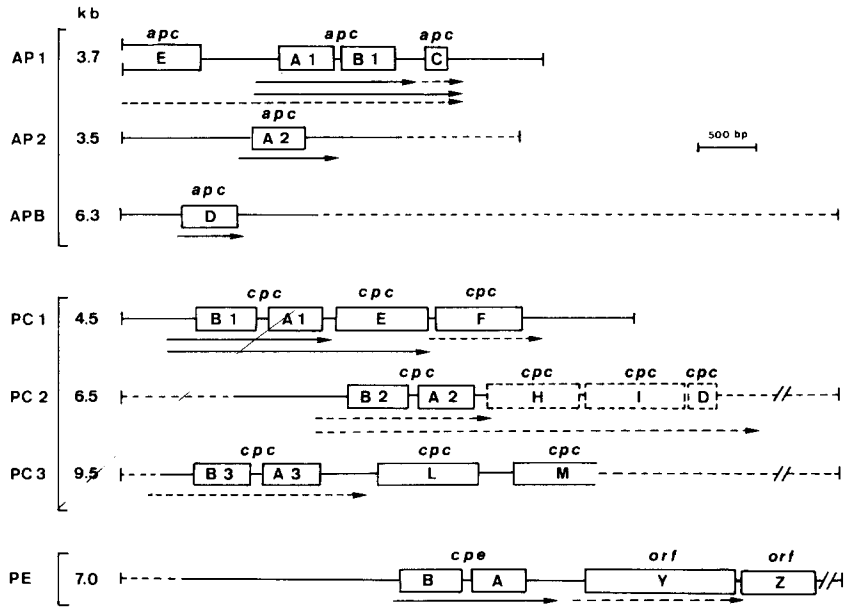

Fig. 2. Physical organization and transcription of genes involved in the biosynthesis of *Calothrix* 7601 phycobilisomes. Each *Eco*R1 fragment is denoted by its size in kilobases (kb). Solid or dotted lines represent totally or partially sequenced DNA, respectively. Solid or dotted arrows represent mRNA mapped or only identified in Northern hybridization experiments, respectively. The different genes are designated as follows: *apcA* and *apcB*, for the α and β subunits of the allophycocyanins (AP1 and AP2); *apcD* for the allophycocyanin B (APB); *cpcA* and *cpcB*, for the α and β subunits of the phycocyanins (PC1, PC2 and PC3); *cpeA* and *cpeB*, for the α and β subunits of the phycoerythrin (PE); *apcE*, for the anchor polypeptide L_{CM}^{92}; *apcC*, for the linker polypeptide $L_C^{7.8}$; *cpcL* and *cpcM*, for rod linker polypeptides; *cpcE, cpcF, orfY* and *orfZ*, for the open reading frames which correspond to unidentified gene products. The genes designated *cpcH, cpcI* and *cpcD* which encode the linker polypeptides L_R^{38}, L_R^{39} and $L_R^{9.7}$, respectively, have been sequenced by Lomax et al. (1987). Some of the previously published gene designations have been changed according to the nomenclature rules recently proposed for cyanobacterial genes (Houmard and Tandeau de Marsac 1988).

of a second 'constitutive' PC species, provisionally called PC3 (Mazel et al.1988). In each PC gene cluster, the *cpcB* gene is located upstream from the *cpcA* gene. The intergenic regions of the *cpc1, cpc2* and *cpc3* gene clusters are 118, 69 and 116 nucleotides long, respectively, and exhibit little or no sequence homology. The deduced amino acid sequences of the *cpcB3* and *cpcA3* genes are 75–78 and 84–85% identical to those deduced from the *cpcB1* or *cpcB2* and from the *cpcA1* or *cpcA2* genes, respectively. Possible roles and location in the phycobilisome structure of the new type of PC (PC3) will be discussed in section 2.3.

Three genes encoding an α-type subunit of AP have been isolated in *Calothrix* 7601, while only two different α-type subunits of APs (αAP1 and

αAPB) have been biochemically characterized in cyanobacterial phycobilisomes. The first gene, *apcA1* (483 bp), is located 70 nucleotides upstream from the gene encoding βAP1, *apcB1* (486 bp). The two other genes, *apcA2* and *apcD*, which are each 483 nucleotides long, are carried by the 3.5 and 6.3 kb *Eco*RI fragments, respectively. No open reading frame which could encode phycobiliprotein subunits has been found in the vicinity of these two genes (Houmard et al. 1988). The deduced amino acid sequence of the predicted *apcD* product is about 70% homologous to the sequence of the APB purified from the cyanobacterium *Synechococcus* 6301 (Suter et al. 1987). It contains the two tryptophan residues specifically found in this phycobiliprotein. The predicted amino acid sequence of the *apcA2* gene shares 59 and 43% identity with those deduced from *apcA1* and *apcD*, respectively (J. Houmard and T. Coursin unpublished data). Additionally, an open reading frame (*apcE*), located upstream from *apcA1*, has been partially sequenced. By analogy to *Nostoc* sp. (Zilinskas et al. 1987), our preliminary data indicate that the putative truncated gene product corresponds to the high molecular weight protein L_{CM}^{92} (J. Houmard and V. Capuano unpublished data).

With the exception of the 3.7, 6.5 and 4.5 kb *Eco*RI fragments which are adjacent on the chromosome (Conley et al. 1986), the other *Eco*RI fragments are not in close proximity to each other (J. Houmard and T. Coursin unpublished data). An alignment of the predicted amino acid sequences of the phycobiliprotein genes is presented in the Appendix (Fig. 11A).

Linker polypeptide genes. Some linker polypeptides have been sequenced from purified proteins (Füglistaller et al. 1984, 1985, 1986). Based on their homology with the deduced amino acid sequences of different open reading frames found in *Calothrix* 7601, three linker polypeptide genes have been identified. The *apcC* gene (204 bp) located 240 nucleotides downstream from *apcB1* encodes the small linker polypeptide $L_C^{7.8}$ associated with AP in the core of the phycobilisome. Its deduced amino acid sequence is 73 and 91% homologous to those from *Synechococcus* 6301 (Houmard et al. 1986) and *Mastigocladus laminosus* (Füglistaller et al. 1984), respectively. The *cpcL* gene (978 bp) and the partially sequenced *cpcM* gene are located downstream from the *cpcA3* gene. These genes encode two polypeptides which are highly homologous to the linker polypeptides L_R^{33} from *Synechococcus* 7002 (D.A. Bryant, cited in Glazer 1987) and $L_R^{34.5}$ from *Mastigocladus laminosus* (Füglistaller et al. 1985, 1986), respectively. Moreover, these genes share approximately 55% identity with the two genes which are located downstream from *cpcA2* and encode the linker polypeptides L_R^{38} and L_R^{39} associated with PC2 in red-light phycobilisomes (Lomax et al. 1987). The *cpcL* and

cpcM genes are thus likely to encode additional linker polypeptides associated with PCs in the phycobilisome rods. A compilation of the predicted amino acid sequences of the linker polypeptides, including those published by Lomax et al. (1987), is presented in the Appendix (Fig. 11B).

Unassigned open reading frames. Four other open reading frames were found by sequence analysis which are likely involved in phycobilisome formation. The first two are *cpcE* (882 bp), located 73 nucleotides downstream from *cpcA1* (Mazel et al. 1988), and *cpcF* (783 bp) which is located 51 nucleotides downstream from *cpcE*. As will be discussed later (see *Regulation of gene expression*), *cpcE* is cotranscribed with the *cpcB1* and *cpcA1* genes in contrast to *cpcF* which is independently transcribed. The two other open reading frames, *orfY* (1290 bp) and *orfZ* (618 bp), have been also completely sequenced. The first one, *orfY*, is located 522 nucleotides downstream from *cpeA*, while *orfZ* is 51 nucleotides downstream from *orfY*. The putative gene products of these four open reading frames do not share significant homology either with known linker polypeptides or with identified phycobiliproteins. Moreover, they could not be assigned to any known proteins by comparison with sequences available in DNA (Los Alamos and EMBL) and protein (NBRF and PseqIP) data banks (D. Mazel unpublished data). Their deduced amino acid sequences are presented in the Appendix (Fig. 11C).

2.2.2 Transcription analysis

The characterization of the mRNA species corresponding to most of the *Calothrix* 7601 genes that we have identified has been carried out by Northern hybridization, S1 nuclease mapping and primer extension experiments using total RNA extracted from cells grown under red- or green-light conditions. The identified mRNA species are presented in Fig. 2.

From this analysis, it appears that only *apcA2, apcD, cpcF* and *orfY* are transcribed as monocistronic units, the sizes of the transcripts being approximately 0.65, 0.55, 1.1 and 1.5 kb, respectively. The 5′ extremities of the *apcA2* and *apcD* transcripts are located 97 and 24 nucleotides respectively upstream from the initiation codon of the corresponding genes. The *apcA1B1C, cpcB1A1E* and *cpcB3A3* gene clusters are transcribed as polycistronic units. The *apcA1B1C* operon is transcribed into three major mRNA species of approximately 1.7 (*apcA1B1C*), 1.4 (*apcA1B1*) and 0.25 kb (*apcC*). The 1.7 and 1.4 kb transcripts have the same 5′ extremity, 193 nucleotides upstream from the initiation codon of the *apcA1* gene. The 5′ end of the shortest transcript has not been mapped, but is likely to be located between the *apcB1* and *apcC* genes. In addition, two transcripts of 5.7 and 5.5 kb, which are about one hundred times less abundant than the 1.4 kb mRNA

species, can also be detected. These correspond to the cotranscription of both the *apcE* gene and the *apcA1B1C* operon. Thus, these large transcripts include most of the genes which encode components of the phycobilisome core. The *cpc1* operon is transcribed as two mRNA species of 1.45 and 2.4 kb corresponding to *cpcB1A1* and *cpcB1A1E*, respectively. These transcripts have the same 5′ extremity which occurs 255 nucleotides upstream from the initiation codon of the *cpcB1* gene. The *cpcB3A3* operon is transcribed as a unique mRNA species of approximately 2 kb. All the genes mentioned above are transcribed at roughly similar levels in cells grown under red or green light conditions and, consequently, are probably independent of the light wavelength control exerted during complementary chromatic adaptation.

The remaining two gene clusters, *cpc2* and *cpe*, which respectively encode the 'inducible' PC2 and PE subunits, are also transcribed as polycistronic units, but only under specific chromatic illumination (Conley et al. 1985, Mazel et al. 1986, our laboratory unpublished data). Two red-light specific transcripts of the *cpcB2A2* operon can be detected. One (1.5 kb) is produced by the cotranscription of *cpcB2* and *cpcA2*. The second transcript (3.8 kb) extends downstream from *cpcA2* and includes the genes encoding the PC2-associated linker polypeptides: L_R^{38}, L_R^{39} and $L_R^{9.7}$ (Lomax et al. 1987, D. Mazel unpublished data). The *cpeBA* operon is transcribed as a unique green-light specific 1.5 kb mRNA species, whose 5′ end occurs 66 nucleotides upstream from the initiation codon of *cpeB* (Mazel et al. 1986, D. Mazel unpublished data).

With the exception of the *cpe* operon in which *E. coli*-like promoter sequences are found, analysis of the nucleotide sequences upstream from the potential start sites of transcription of the different mRNA species identified in *Calothrix* 7601, did not reveal sequences similar to the *E. coli* consensus promoter sequences (Fig. 3A). Moreover, no consensus sequence which might correspond to the binding site for the *Calothrix* 7601 RNA polymerase could be reasonably deduced from the comparison of all the sequences which are available today (Fig. 3A). In contrast, a striking homology (56%) is found between a 43 bp long sequence (located 41 nucleotides upstream of and including the first two nucleotides of the *Calothrix* 7601 *cpc1* transcripts) and a sequence of *Synechococcus* 7002 (332 to 374 nucleotides upstream from the AUG of the *cpcB* gene), which, according to Pilot and Fox (1984), probably corresponds to the promoter region (Fig. 3B). Similarly, in *Anabaena* 7120, the nucleotide sequence (nucleotides −287 to −243 in Belknap and Haselkorn 1987) which precedes and includes the transcription start site (nucleotide −243) is 73% homologous to that of *Calothrix* 7601 (Fig. 3B). In contrast, no other homology has been noticed

Fig. 3. Nucleotide sequences of the promotor regions. A. Comparison of the sequences from the *Calothrix* 7601 genes. Distances from the first ATG of the coding sequences are indicated in base pairs (bp). Numbering of nucleotides is based on the assignment of the first base of the transcripts as nucleotide + 1. B. Comparison of the sequence of the promoter regions of the *Calothrix* 7601 *cpcB1A1E* operon with those of the *cpc* operon from *Synechococcus* 7002 (nucleotides 1 to 59 in De Lorimier et al. 1984) and from *Anabaena* 7120 (nucleotides –294 to –236 in Belknap and Haselkorn 1987). Lower cases correspond to the start sites of transcription determined for *Calothrix* 7601 (also indicated by an arrow) and for *Anabaena* 7120. Stars indicate nucleotide identities. Dashes indicate gaps inserted to maximize homology. Numbering of nucleotides is based on the assignment of the first base of the *Calothrix* 7601 *cpc* transcripts as nucleotide + 1.

within the 5′ untranslated regions of the transcripts. It is thus tempting to hypothesize that the homologous part of these sequences plays a role in initiating transcription of the PC genes in these cyanobacterial strains (Mazel et al. 1988).

The 3′ end of the *cpcB1A1* transcript has also been mapped. Transcripts terminate a few nucleotides downstream from a thermodynamically stable stem-and-loop structure which is not followed by an AU-rich sequence (Mazel et al. 1988). This result is similar to those obtained for the 3′ ends of the transcripts of the *gvpA1A2C* operon (see *Transcription of the gvp genes)*.

2.2.3 Regulation of gene expression

Phycobilisomes are supramolecular structures in which the stoichiometry of the different components and their positioning within the structure are essential for optimizing their light-harvesting capacity and their ability to transfer energy to the reaction centers of photosystem II. Moreover, the energy trapped by the phycobilisomal units generally does not exceed the functional capacity of the photosynthetic reaction centers. To do this, some

cyanobacteria are not only able to alter the number and the size of their phycobilisomes in response to the incident light energy, but also to modify their rod composition in response to light wavelength. These properties rely on an accurate photoperception system coupled to precise coordination of the expression of the different genes involved in these light-dependent processes. At the molecular level, this can be performed by varying the copy number and/or by regulatory mechanisms which can operate at different levels: transcriptional (level of transcription and/or mRNA turnover), translational (occurrence or modulation of the protein synthesis) and post-translational (maturation and/or degradation of proteins).

Based on physiological studies performed on exponentially growing cells, it has been established that phycobiliproteins are stable proteins which do not turnover (Bennett and Bogorad 1973, Tandeau de Marsac 1977). After transfer from red to green illumination, or *vice versa*, changes in the phycobiliprotein content of the cells are due to *de novo* synthesis, the preexisting phycobiliproteins disappearing only by dilution during growth. Experiments performed using transcription inhibitors indicated that adaptation to new illumination conditions were probably due to transcriptional regulation of the expression of the genes encoding PE and most probably PC2 (Gendel et al. 1979, Tandeau de Marsac 1983).

With the recent characterization of the genes encoding components of the phycobilisomes in *Calothrix* 7601 and the analysis of their transcription using total RNA from cells grown under red- or green-light conditions, it has been definitely established that the transcription of the *cpcB2A2HID* and *cpeBA* operons is red- and green-light specific, respectively (Fig. 2; Conley et al. 1985, Mazel et al. 1986, our laboratory unpublished data). However, this regulatory mode constitutes only part of the complex transcriptional regulatory system which occurs in *Calothrix* 7601. Indeed, at least three operons, *cpcB1A1E, cpcB2A2HID* and *apcEA1B1C*, give rise to segmented transcripts which are present at different levels. It is also worth noting that in each of these three operons, and also in the *cpeBA* and *cpcB3A3* operons, the genes coding for the α and β subunits of a specific phycobiliprotein are always cotranscribed, ensuring that both subunits can be made in an equimolar ratio. Besides the *cpcB3A3* operon and the *apcA2* gene, which are transcribed about 50 times less efficiently than the *cpcB1A1* and the *apcA1B1* genes, respectively, and to which no function has yet been assigned, we generally observe a good correlation between the levels of transcription of the different operons and the proportions of the corresponding phycobiliproteins in the phycobilisome. This suggests that the promoters of these operons may have different efficiencies and/or require specific effectors. The occurrence of segmented transcripts for the

apcEA1B1C and *cpcB2A2HID* operons, and the higher level of the mRNA species which do not contain the linker sequences, constitute another means of coordinating the expression of genes whose products are required in different molar ratios (Grossman et al. 1986, J. Houmard and V. Capuano unpublished data). The occurrence of a small mRNA species (0.25 kb) corresponding to the *apcC* gene which encodes the AP-specific linker $L_C^{7.8}$, is somewhat more surprising. As also noticed by Grossman et al. (1986), we observe that this small transcript is more abundant in cells grown under red light than under green light, which is unexpected given the invariance of the core composition during complementary chromatic adaptation. Conversely, the 1.7 kb mRNA (*apcA1B1C*) is less abundant in cells grown under red light than under green light, while the amount of the 1.4 kb species (*apcA1B1*) is almost unchanged. Since the sum of the 0.25 kb and 1.7 kb transcripts is about constant under both light regimes, this suggests that the small transcript might result from the processing of the longer one.

In *Synechococcus* 7002 (Bryant et al. 1986), the *cpc* gene cluster consists of six adjacent genes, namely *cpcB* (βPC), *cpcA* (αPC), *cpcC* (L_R^{33}) and *cpcD* ($L_R^{9.7}$), followed by two open reading frames, *cpcE* and *cpcF*. In *Anabaena* 7120 (Belknap and Haselkorn 1987), at least five genes have been found which are organized similarly and are highly homologous to the *cpcBACDE* genes from *Synechococcus* 7002. Surprisingly, the *Calothrix* 7601 *cpcB1* and *cpcA1* genes are followed by two open reading frames whose deduced amino acid sequences bear no significant homology with any of the known linker polypeptide sequences. However, the first open reading frame, *cpcE*, which is cotranscribed with the *cpcB1* and *cpcA1* genes, shares approximately 65% identity with the *cpcE* gene located immediately downstream from *cpcD* in *Synechococcus* 7002 and *Anabaena* 7120 (Belknap and Haselkorn 1987, D.A. Bryant personal communication). Finally, unlike *Synechococcus* 7002 in which there is a short open reading frame (114 bp long) within the transcript upstream from the *cpcB* gene (De Lorimier et al. 1984, Pilot and Fox 1984), no open reading frame has been found upstream from *cpcB1* in *Calothrix* 7601, although this untranslated region is also rather long (255 bp).

As shown in Fig. 4, an interesting feature of the *cpcB1A1E* operon is that 22 nucleotides (sequence 1), located at the 5′ extremity of the transcript, are able to form a stem-and-loop structure ($\Delta G = -14.6$ kcal/mol) with the 22 nucleotides (sequence 3) located 35 nucleotides downstream. In addition, 11 nucleotides (sequence 2), located inside the loop of this secondary structure, are able to form a stem-and-loop structure ($\Delta G = -9.8$ kcal/mol) with part of sequence 1 (sequence 1′) which is involved in the formation of the stem of the structure 1–3. In contrast to structure 1′–2, the larger and

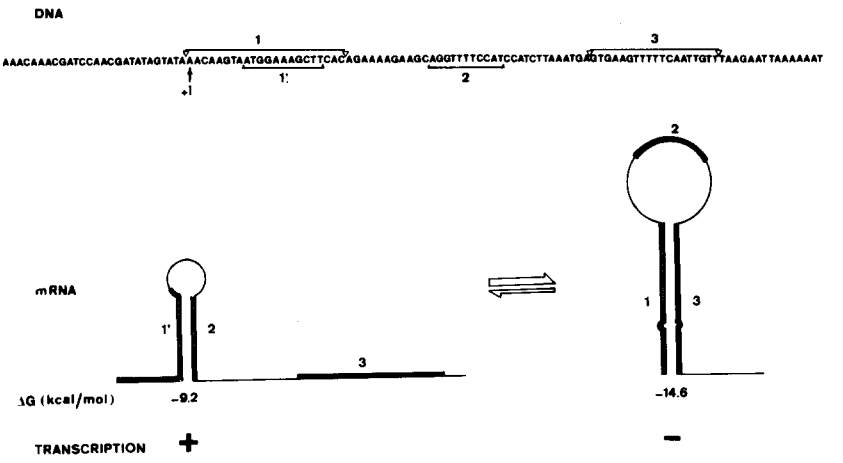

Fig. 4. Hypothetical attenuation-like regulatory mechanism for the transcription of the *Calothrix* 7601 *cpcB1A1E* operon. The arrow denotes the start site of transcription (+1). Numbers indicate the DNA sequences which are involved in the formation of secondary structures. The upper line coresponds to the DNA sequence and the bottom of the figure presents the equilibrium between the two potential secondary structures which can be formed in the 5′ untranslated region of the mRNAs. The free energy values (ΔG) were calculated according to Cech et al. (1983).

more stable one (structure 1–3) is followed by a very AT rich region. These observations, together with the length of the 5′ untranslated region of the mRNA, lead us to propose an 'attenuation' model, in which structure 1–3 could act as an early transcription terminator, while structure 1′–2 would allow transcription to proceed. A positive effector could stabilize structure 1′–2, as soon as the corresponding sequences are transcribed, in order to prevent the formation of the potential 'terminator' (structure 1–3). In this hypothetical model, the equilibrium between these two potential structures would modulate the transcription rate of the *cpcB1A1E* operon in response to environmental factors, such as light intensity or nutrient availability (Mazel et al. 1988).

On the other hand, while the *cpcB1A1* and *cpcB1A1E* transcripts have the same 5′ extremity, they are present in different amounts. Thus, there must be a difference either in the termination efficiency or in the stability of these two mRNAs. Potential secondary structures in the mRNAs have been postulated to play an important role in transcription termination in prokaryotes (Platt 1986). However, recent results suggest that stem-and-loop structures might be important in protecting against 3′ exoribonucleases, rather than in terminating transcription (Brawerman 1987). Indeed, it has been clearly demonstrated that these structures can influence the rate of mRNA

decay and increase their stability by preventing degradation (Belasco et al. 1985, Wong and Chang 1986, Newbury et al. 1987). In the case of the *cpcB1A1* transcript, an extremely stable stem-and-loop structure ($\Delta G = -30$ kcal/mol) could be formed ending two nucleotides upstream from its 3′ end. In contrast, no such stable secondary structure exists downstream from the *cpcE* stop codon and, on Northern blots, the band corresponding to the *cpcB1A1E* mRNA is more diffuse than that of *cpcB1A1* mRNA, suggesting a lower stability of the *cpcB1A1E* mRNA. Although we cannot rule out that these hairpin structures play a role in the release of the RNA polymerase, it is tempting to correlate their presence with the stability and abundance of the two mRNA species corresponding to *cpcB1A1* and *cpcB1A1E*.

Finally, the fact that the *cpcE* gene is cotranscribed with the *cpcB1* and *cpcA1* genes but does not encode a phycobilisomal protein-like product, raises the questions concerning the function of the *cpcE* gene product. Among other possibilities, it might encode a factor involved in a post-translational process, such as the linkage of the chromophore to the apoproteins or the maturation of the polypeptides. Examples of post-translational modifications of the phycobilisome components have recently been described, including methylation of the phycobiliprotein β subunits (Minami et al. 1985, Klotz et al. 1986, Rümbeli et al. 1987), reversible phosphorylation (Allen et al. 1985) and glycosylation of some linker polypeptides (Riethman et al. 1987). Some of these modifications might be necessary to prevent proteolytic degradation of the phycobilisome components and/or to confer a functional configuration. Analysis of a mutant obtained by interposon mutagenesis (Bryant 1988) indicated that the *cpcE* gene is probably implicated in chromophore attachment to the PC α subunit in *Synechococcus* 7002. This would be in agreement with our first hypothesis. However, confirmation of this result awaits further transcription analysis in order to rule out a possible polar effect of the mutation studied on the expression of the *cpcF* gene located downstream from the inactivated *cpcE* gene in *Synechococcus* 7002.

In *Calothrix* 7601, the *cpcF* gene found downstream from *cpcE*, is transcribed as a unique mRNA species and could have thus been considered as a gene unrelated to the formation of phycobilisomes. However, analysis of spontaneous pigmentation mutants from *Calothrix* 7601 strongly suggests that the *cpcF* gene product is involved in some regulatory processes in the synthesis of phycobilisomes. Indeed, the characterization of the pigmentation mutant, GY3, revealed that a typical bacterial insertion sequence, IS*701*, had spontaneously inserted in the 5′ part of *cpcF*, leading to transcription termination. Phycobilisomes purified from this mutant are smaller than those from the wild-type strain. They are practically devoid of PCs and

PE and, consequently, do not contain rods, while linker polypeptides appear to be present. Although red-light phycobilisomes from this mutant contain less than 10% of PC2 when compared with the wild-type phycobilisomes and green-light phycobilisomes contain only traces of PE, the synthesis of these phycobiliproteins remains under the chromatic light control (D. Mazel unpublished data).

Even more interesting are the results of the anlaysis of the transcription of the phycobiliprotein genes in the mutant GY3. They revealed that the *apcA1B1*, *cpcB1A1* and *cpcB2A2* operons are transcribed to the same extent as in the wild-type strain, while the transcription of the *cpeBA* operon is dramatically decreased. However, no free PCs (or apoproteins) have been detected by biochemical and immunological techniques although the AP1 concentration is unmodified. This suggests that the *cpc* genes are transcribed but are not translated or are only poorly translated. One cannot totally exclude, however, that the apoproteins are very rapidly degraded. Since a single gene product is unlikely to be directly involved in the translational regulation of one operon and in the transcriptional regulation of another one, we propose, as a working hypothesis, that the *cpcF* gene product is primarily acting on the translation of the *cpc* genes, the effect on the transcription of the *cpe* genes being a secondary consequence of this translational regulation. Indeed, since PE hexamers cannot attach directly to the phycobilisome core, an autoregulation of the transcription of the *cpe* genes by free PE could occur. (D. Mazel unpublished data).

If our hypothesis is correct, how could the *cpcF* gene product act on the translation of the *cpc* genes? Potential secondary structures in the long 5′ untranslated region of the *cpcB1A1* operon and in the region between *cpcB1* and *cpcA1* might mask the ribosome binding site sequences. We thus propose a model of regulation (Fig. 5) in which the *cpcF* gene product acts as a positive effector destabilizing the hairpin structures upstream from *cpcB1* and allowing the ribosome to bind. Subsequently, the migration of the ribosomes along the mRNA would liberate the ribosome binding site upstream from the *cpcA1* gene allowing translation of this gene to proceed. Such a translational regulatory mode, which is independent of the light wavelength, would permit rapid changes in the synthesis of PC1 in response to environmental factors. Since hairpin structures are also present in the same region of the transcripts for the *cpcB2A2* and *cpcB3A3* operons, this suggests that these operons could also be regulated by a similar mechanism.

2.3 Conclusions

With the exception of the gene encoding one of the phycobiliproteins of the phycobilisome core, β18.3, all of the *Calothrix* 7601 phycobiliprotein genes

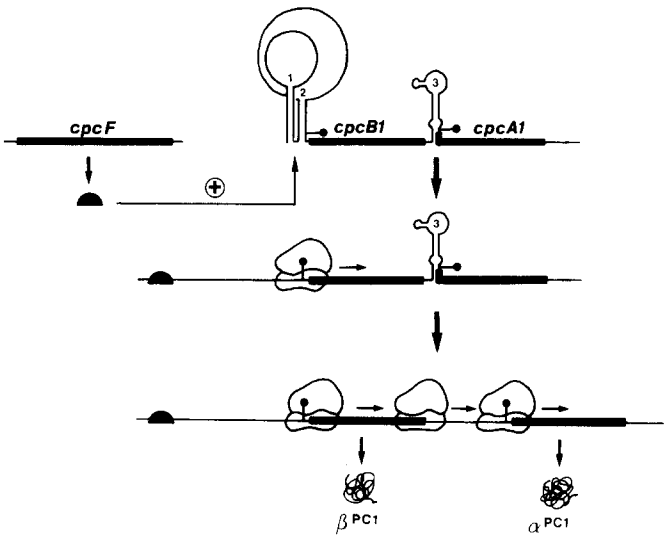

Fig. 5. Hypothetical model for post-transcriptional regulation of the *cpcB1A1E* operon mediated by the *cpcF* gene product. Boxed regions denote the coding regions of the genes. Hairpins 1, 2 and 3 are the potential secondary structures which might form within the *cpcB1A1* and *cpcB1A1E* transcripts (free energy values $\Delta G = -14.6$, -7.5 and -22.3 kcal/mol, respectively, calculated according to Cech et al. 1983). ↑ indicates the mRNA sequences to which ribosomes bind to initiate translation. The filled half-circle represents the *cpcF* gene product.

have now been identified, as have most of the genes encoding linker polypeptides. At least two levels of control are involved in the regulation of the expression of these genes: transcriptional and translational. The first mode of control occurs in response to changes in the chromatic light, but, according to our proposed 'attenuation' model, it might also occur in response to other environmental changes. The translational mode appears to be independent of the chromatic illumination, and is likely to be a more general process involved in modulating the synthesis of phycobilisome components.

Three sets of genes encoding both the α and β subunits of PCs and three genes encoding α-type subunits of AP have been found in *Calothrix* 7601. The functions of *cpcB3A3* and of *apcA2* remain to be determined. If their gene products are accommodated within the phycobilisome structure, the divergence of their amino acid sequences, as compared with those of the other *cpc* and *apc* gene products, suggests that they may have acquired different roles and/or lead to heterogeneity in the phycobilisome building blocks. Alternatively, these phycobiliprotein gene products might be specific for different cellular types. This second hypothesis can probably be ruled out, since results from experiments performed with *Calothrix* 7601 and 7504 (which, in contrast to *Calothrix* 7601, is still able to differentiate fully

functional heterocysts) have shown that the level of transcription of the *cpcB3A3* genes is similar both in heterocysts and in vegetative cells (D. Mazel unpublished data). In contrast, detailed analyses of the structure and composition of the phycobilisome core from *Synechococcus* 6301 and *Synechocystis* 6701 have shown that the constituent trimers are heterogeneous. This heterogeneity is correlated with the function of the different phycobiliproteins in the transfer of energy to the photosynthetic reaction centers (Glazer 1982, Anderson and Eiserling 1986). It might well be that some of the AP α-type subunits found in the core correspond to the *Calothrix* 7601 *apcA2* gene product. Similarly, although the only example of rod heterogeneity reported so far is that of the PE- or PC-rich rods present in phycobilisomes from *Nostoc* sp. MAC (Glick and Zilinskas 1982, Anderson et al. 1983), heterogeneity of the 'constitutive' PC molecules could exist either within different hexamers or within one hexamer of the peripheral rods. In fact, this hypothesis would be in full agreement with the results obtained by G. Guglielmi and D.A. Bryant (unpublished data). These authors observed that PE-rich rods from *Nostoc* 8009 phycobilisomes purified from white-light grown cells only contain one PC hexamer which is composed of three types of PCs (two 'constitutive' PCs and one 'inducible' PC).

Several other genes, besides the genes encoding phycobilisome components, are required for the synthesis of the chromophores and their linkage to the apoproteins. Regulatory genes might also be necessary to ensure coordinated synthesis of the chromophore(s) and of the apoprotein. Furthermore, in the case of chromatically adapting strains, several other genes are expected to be involved in the synthesis of the photoreceptor. Particular efforts should now be directed towards their identification. The study of the function of the *apcA2* gene product is of particular interest with regards to that problem. Indeed, we cannot exclude that the *apcA2* gene product corresponds to the photoresponsive AP found in *Calothrix* 7601 by Ohad et al. (1979, 1980). Nevertheless, the question of the relationship between such a photoresponsive AP and the photoreversible pigment involved in complementary chromatic adaptation remains to be solved.

3. Hormogonia differentiation

As shown in Fig. 6 (A and B), vegetative filaments of *Calothrix* 7601 are morphologically different from differentiated hormogonia. Hormogonia cells are smaller and filled with refringent granules which are aggregates of gas vesicles providing cells with buoyancy. Pili are produced at the

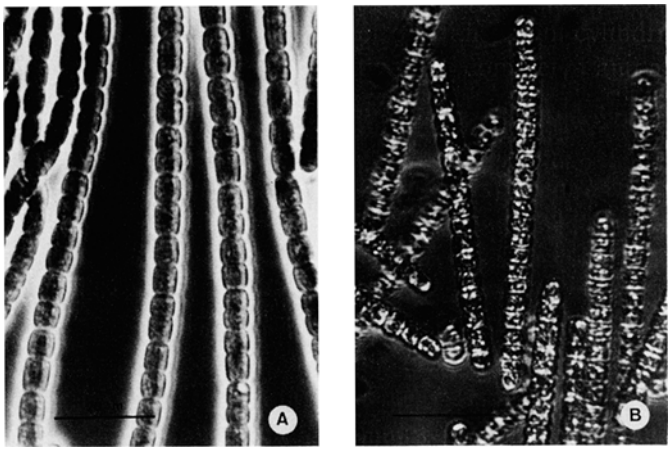

Fig. 6. Light micrographs of *Calothrix* 7601 vegetative trichomes (A) and differentiated hormogonia (B). Bar marker indicates 20 μm.

cell surface. Their presence is correlated with hormogonia motility (G. Guglielmi unpublished data).

3.1 Physiological characteristics of hormogonia differentiation

Physiological studies of hormogonia differentiation in *Calothrix* 7601 revealed that it may begin in any cell within the filament and then progresses to neighboring cells along the filament. Moreover, the percentage of differentiated cells has been shown to be highly dependent on i) the spectral quality of the light, ii) the growth phase of the preculture and iii) the incubation temperature. Indeed, 90 to 100% differentiation is obtained if cells in early exponential phase are transferred to red illumination and incubated at 25–30 °C. If, under similar conditions, cells are incubated either under green light or under 50% red light plus 50% green light, instead of red light alone, the percentage of differentiation decreases to less than 30% or to approximately 50%, respectively. Interestingly the efficiency of differentiation is lower for cells in mid- or late-exponential phase. Indeed, 90–100% differentiation is only obtained if such cells are transferred into fresh culture medium and then incubated under red light (T. Damerval and G. Guglielmi unpublished data).

Thin sections of cells collected at different times after the induction of the differentiation (Fig. 7) show that cells collected 6 h after the induction are dividing with a reduced elongation of the longitudinal cell wall. They produce gas vesicles which appear as electron-transparent cylindrical structures. After 12 h, cell divisions (usually two) are completed and aggregates

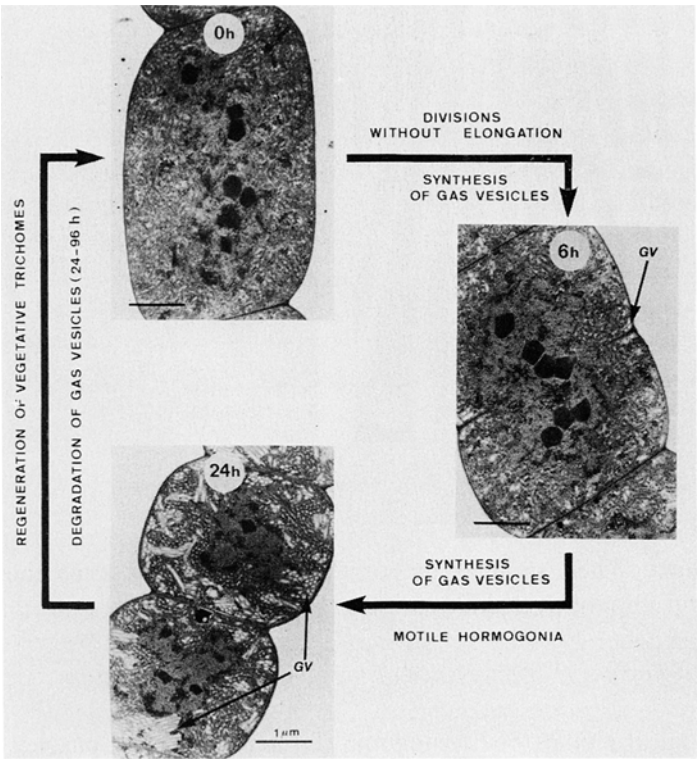

Fig. 7. Developmental cycle of hormogonia differentiation in *Calothrix* 7601. Electron micrographs of thin section of cells collected at various times (h) after the induction of the differentiation. Gas vesicles (GV) are visible as electron transparent structures. Bar marker indicates 1 μm.

of gas vesicles increase in size until 24 h. This hormogonia state is transient; as growth proceeds, regeneration of vegetative filaments occurs with a concomittant loss of the gas vesicles (T. Damerval and G. Guglielmi unpublished data). Figure 8 presents an electron micrograph of a gas vesicle isolated from *Calothrix* 7601.

3.2 Functional organization of the genes involved in gas vesicle formation

In order to elucidate the molecular basis of hormogonia differentiation, we looked for an appropriate marker. Since the formation of gas vesicles appears to be correlated with the other morphological changes which occur during hormogonia differentiation, we chose to isolate the gene encoding the major structural protein (GVP) of the gas vesicle.

Fig. 8. Electron micrograph of a negatively stained gas vesicle isolated from *Calothrix* 7601.

3.2.1 Characterization of the gvp genes

Gas vesicle structure and function have been analyzed in detail (for reviews see Walsby 1978, 1981, 1987). According to studies by Walsby and coworkers, these vesicles, which are impermeable to water and permeable to gases, are made up of only one protein species whose amino acid sequence is highly conserved among cyanobacteria and other gas-vacuolate bacteria (Walker et al. 1984). It was thus possible to design a synthetic 29 nucleotide-long oligodeoxynucleotide sequence in order to probe the *Calothrix* 7601 genome. Two *Hin*dIII fragments of 2.6 and 2.3 kb hybridized with this probe. Both fragments were cloned from a genomic library constructed in λEMBL3 and the genes carried by these two DNA fragments were further analyzed. The 2.6 kb fragment bears two genes encoding the structural gas-vesicle protein. These genes, *gvpA1* and *gvpA2* are identical in length (216 bp) and are separated by 105 nucleotides. Their nucleotide sequences are 91.5% homologous. The 18 nucleotide differences are scattered along the length of the genes, but generally affect only the third base of a codon (Tandeau de Marsac et al. 1985, Damerval et al. 1987). These genes most likely arose by duplication of an ancestral gene followed by silent mutations. In contrast, the upstream and downstream regions of each of these two copies share no homology. Each of these two genes encode a protein of 7375 daltons (see Appendix, Fig. 12). The first methionine residue is apparently proteolytically processed (Tandeau de Marsac et al. 1985), since it has not been found in the amino acid sequence performed on the purified protein (Hayes et al. 1986). As expected from biochemical studies (Walsby 1978, Walker et al. 1984), the *gvpA* gene product is highly hydrophobic and shares a very high degree of homology with the total or partial amino acid sequences determined for the gas vesicles protein from other cyanobacteria and

from halobacteria (Tandeau de Marsac et al. 1985, Damerval et al.1987, Das Sarma et al. 1987).

A third open reading frame has been found downstream from the *gvpA2* gene. It could not be correlated to any known gene or product, but has been called *gvpC* since it is cotranscribed with *gvpA1* and *gvpA2* (see Section 3.2.2). A potential ribosome binding site, GGAG, located 6 nucleotides upstream from an ATG sequence suggests that the *gvpC* gene starts 215 bp downstream from the end of *gvpA2* and would be 489 bp long. Unlike the gas vesicle protein, the *gvpC* gene product is predominantly hydrophilic. Interestingly, the internal part of the *gvpC* gene is composed of four contiguous repeats, each 99 bp long, which form highly homologous repeats in the deduced amino acid sequence (see Appendix, Fig. 12). Another kind of periodicity has been detected inside the 99 bp repeats, suggesting that the *gvpC* gene might have evolved by amplification of a 33 bp long primordial building block (Damerval et al. 1987).

A fourth gene involved in gas vesicle formation is present in the *Calothrix* 7601 genome. Indeed, the 2.3 kb fragment which hybridized with the 29-mer oligonucleotide also hybridized with a probe corresponding to an internal region of the *gvpA1* gene. This gene, designated *gvpD*, is located approximately 5 kb upstream from the other *gvp* genes. Preliminary sequence data indicate that *gvpD* is 252 nucleotides long and could encode a protein of approximately 8600 daltons. Consequently, the *gvpD* gene product is, like the *Halobacterium halobium* GVP (Das Sarma et al. 1987), a few amino acids longer than that of the *gvpA* genes. The sequence of the N-terminal 70 amino acid residues is about 80% identical to that of the *gvpA* gene products (T. Damerval unpublished data).

3.2.2 Transcription of the gvp genes

As expected for a differentiation process, the *gvp* genes are transcribed specifically in hormogonia and not in vegetative cells. For the purpose of a detailed transcriptional analysis of the different mRNA species corresponding to the *gvp* genes, total RNA was extracted from hormogonia 6 h after the induction of differentiation (see section 3.3). Hybridization experiments, coupled with mapping of the 5′ and 3′ ends of the transcripts, demonstrated that the 0.3, 0.8 and 1.4 kb long mRNA species correspond to the transcription of *gvpA1*, *gvpA1A2* and *gvpA1A2C*, respectively. The fourth mRNA species, 0.6 kb long, corresponds to the transcription of *gvpD* (Csiszàr et al. 1987). Thus, only the first three genes are organized in an operon which is segmentally transcribed. As shown in Fig. 9, these transcripts have the same 5′ extremity, but different 3′ ends located, in each case, a few nucleotides downstream from stem-and-loop structures, the thermodynamic stability of

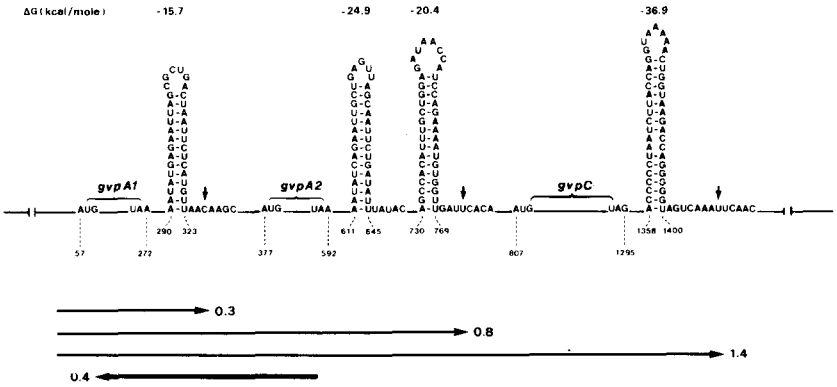

Fig. 9. Potential secondary structures in relation to the transcription of the *gvpA1A2C* operon. Vertical arrows indicate the 3′ ends of the transcripts. Numbering of nucleotides is based on the assignment of the first base of the *gvpA1A2C* operon transcript as nucleotide + 1. The horizontal arrows depict the three mRNA species which start at nucleotide + 1, as well as the antisense RNA which starts at nucleotide 576. The sizes of the transcripts are indicated in kilobases. The free energy values (ΔG) of the potential secondary structures were calculated according to Cech et al. (1983). The *gvpA2* gene was previously designated *gvpB* (Damerval et al. 1987).

which increases gradually towards the end of the operon. The start site of transcription of the *gvpA1A2C* operon is located 56 nucleotides upstream from the initiation codon of the *gvpA1* gene. *E. coli*-like promotor sequences are present upstream from this start site of transcription (TAGTAT as '–10' and TTTACA as '–35'; Fig. 3A). Interestingly, we also found an antisense RNA, approximately 0.4 kb long, which starts in *gvpA2* and ends in *gvpA1*. This antisense RNA can form a perfectly matched duplex with the three mRNA species corresponding to the *gvpA1A2C* operon. No sequences sharing homology with either the *gvpA1A2C* or the *E. coli* promotor sequences were found upstream from its mapped transcription start site (Fig. 3A; Csiszàr et al. 1987).

3.3 Regulation of gene expression during hormogonia differentiation

Hormogonia differentiation is a complex developmental process which leads to rapid changes in the metabolic state of the vegetative cells and probably involves coordinated regulatory mechanisms. In a first attempt to elucidate these mechanisms, it was of interest to examine the kinetics of transcription of the genes involved in the gas vesicle formation, as well as those involved in the synthesis of the phycobiliproteins after the induction of and during this differentiation process. For this purpose, total RNA extracted from

cells, collected at different times after the induction of the differentiation, was hybridized with probes specific for either the gas vesicle (*gvpA1*) or the phycobiliprotein (*cpcB1A1* and *apcA1B1*) genes. The latter two probes were chosen because the corresponding genes are expressed in vegetative cells irrespective of the light conditions (see section 2.2.2). Similar results were obtained with these two probes. Figure 10 presents the results of this kinetic study. At the time of the transfer (0 h), no transcript is revealed with the *gvpA1* probe. Transcription of the *gvp* genes and of the antisense RNA becomes maximal between 3 and 6 h later. The preexisting transcripts are then rapidly degraded (Fig. 10). After 9 to 12 h, probes internal to either the *gvpA1* or the *gvpC* genes revealed the same smear of short mRNA species (data not shown). This indicates that mRNA degradation probably results from both endonucleolytic cleavages and 3′ exoribonucleases activities. In contrast, the amount of the antisense RNA apepars more constant over the same time period. This suggests that there are differences in the kinetics of appearance and/or stability of these two types of RNAs. Conversely, the transcription of genes encoding phycobiliproteins is totally and specifically arrested only when the genes involved in the formation of the gas vesicles are transcribed (Fig. 10). These results clearly demonstrate that the expression of both the *gvp* and the phycobiliprotein genes is regulated at the level of transcription during hormogonia differentiation (T. Damerval and G. Guglielmi unpublished data). Some of the genes are turned on while others are turned off. Our results suggest that different RNA polymerase sigma

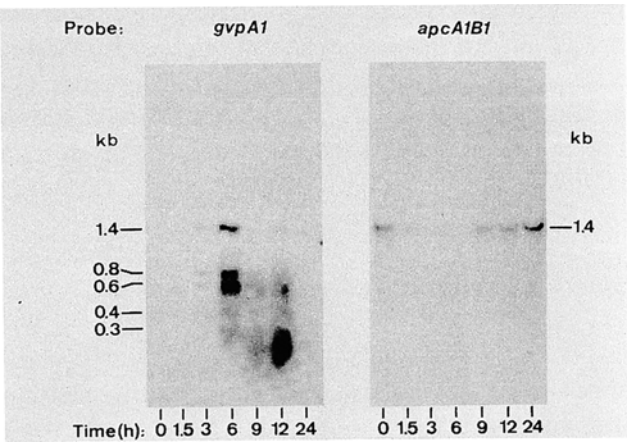

Fig. 10. Kinetics of the transcription of the gas vesicle protein (*gvp*) and of the allophycocyanin (*apc*) genes in *Calothrix* 7601. Total RNA was extracted at various times (h) after the induction of hormogonia differentiation and hybridized with probes internal to the *gvpA1* and *apcA1B1* genes. The size of the transcripts is indicated in kilobases (kb).

factors may exist allowing transcription to occur at the different developmental stages. Moreover, the presence of an antisense RNA, together wtih the rapid degradation of the transcripts corresponding to the *gvp* genes, indicate that the expression of the *gvpA1A2C* operon is also regulated at a post-transcriptional level (Csiszàr et al. 1987).

2.4 Conclusions

Until recently, it was thought that gas vesicle formation results from an autocatalytic phenomenon involving a single protein species (Jost and Jones 1970, Walsby 1981). We have shown that four genes are involved in gas vesicle formation in the cyanobacterium *Calothrix* 7601. The two structural genes, *gvpA1* and *gvpA2*, specify identical products despite the mutations which have occurred within the coding sequences. The very small size and very particular function of this protein may allow only few modifications to occur in its primary structure without altering the function of the molecule. In addition, the fact that this gene has been duplicated might be related to the very high level of expression required during the differentiation process. Indeed, large amounts of protein must be synthesized within a very short time. Multigene families are usually found whenever expression must be faster than that which can be achieved by enhanced expression from a single copy (Stark and Wahl 1984). On the other hand, the existence of the *gvpD* gene, which is very homologous to *gvpA* indicates that two protein species are most probably required to make the two different parts of the gas vesicle: the cylindrical body and the conical ends (Fig. 8).

Another unexpected result revealed by our studies concerns the discovery of the *gvpC* gene. Although its product has not yet been characterized biochemically, there are indications that this gene is translated in *Calothrix* 7601. Essentially two roles can be anticipated for this putative gene product: binding of the gas vesicle protein so as to maintain it in a soluble form suitable for assembly, or regulation of the gas vesicle formation. The primary structure of the *gvpC* gene product and the results of a systematic search for the presence of the *gvpC* gene in different cyanobacterial strains which produce gas vesicles either 'constitutively' or after induction, lead us to favor the first hypothesis (T. Damerval and A.-M. Castets unpublished data). Indeed, the role of this protein could be to increase the efficiency of the gas vesicle assembly, the combination of the gas vesicle proteins with this binding protein being a dispensable step in the assembly process.

Finally, we have shown that the *gvp* genes are rapidly turned on during hormogonia differentiation, while those encoding phycobiliproteins are simultaneously and temporarily turned off. The antisense RNA could form

double-stranded duplexes with part of the *gvpA1A2C* transcripts, thereby favoring their rapid degradation by double-stranded RNA-specific endoribonucleases. The same antisense RNA could also block the translation of the *gvpA1* and *gvpA2* genes by impeding the migration of the ribosomes. These results constitute the first evidence that this complex differentiation process is regulated at both transcriptional and post-transcriptional levels.

4. Concluding remarks

Molecular studies of *Calothrix* 7601 genes involved in phycobilisome synthesis and hormogonia differentiation have yielded important new information concerning their physical organization and the regulation of their expression. However, major questions related to these photoregulated processes remain unanswered. The most intriguing ones are: how does light act at the molecular level; does the same photoreceptor trigger the modulation of gene expression or do two or more systems coexist in this cyanobacterium; are different effectors (sigma factors?) involved in these two processes? At present speculations, but no definite answers, can be formulated. Undoubtedly, light wavelength plays an important role in both processes. However, it is worth noting that there are differences in the way cells respond to the chromatic light stimuli. In the case of complementary chromatic adaptation, PC2 is synthesized under red illumination but also in darkness, while PE synthesis can only occur under green illumination (Tandeau de Marsac 1983). Furthermore, this antagonistic effect of red versus green illumination is efficient and fully reversible at any stage of exponential cell growth. Red light also stimulates hormogonia differentiation, but, in contrast to chromatic adaptation, hormogonia differentiation does not occur in the dark under our experimental conditions, i.e. when the preculture is grown under white light. Moreover, as growth of differentiated hormogonia proceeds in red light, regeneration of vegetative filaments occurs without any green light requirement. Hormogonia differentiation thus appears to be rather similar to the way in which red light triggers a cascade of events, which lead to the final biological response in higher plants, such as the stimulation of seed germination (Cone and Kendrick 1986).

In fact the differences between the two systems, phycobilisome biosynthesis and hormogonia differentiation, reinforce the interest in their study in the same organism. Recent and extremely important progress has been made by Wolk and coworkers (Wolk et al. 1984, Flores and Wolk 1985) in establishing a conjugation system for filamentous cyanobacteria. This genet-

ic transfer system has been adapted to *Calothrix* 7601 by J.G. Cobley (personal communication). In addition, numerous mutants are now available. With these genetic tools, we expect to obtain deeper insights into these complex but fascinating photoregulatory processes in the near future.

Acknowledgements

We are very grateful to Dr G. Cohen-Bazire for valuable comments and support, to Dr D.A. Bryant for his very fruitful collaboration in this research project, to K. Csiszàr for her valuable contribution to this work and to T. Coursin and A.-M. Castets for their skillful technical assistance. We would like also to thank Dr A. Pugsley for a critical reading of the manuscript. The work presented in this review was supported by ATP-CNRS grant (Microbiologie 955353, 1984) and AI-CNRS 990019.

Appendix

Access to the sequences in the EMBL sequence data library

Several of the nucleotide sequences of the *Calothrix* 7601 genes which are mentioned in this paper are or will soon be available in the EMBL/GenBank under the following mnemonics:

FDAPCD	- Y00539	for *apcD*,
FDCPCABE	- X06084	for *cpcB1A1E*,
FDCPCBA2	- X06451	for *cpcB2A2*
FDCPCAB3	- X06083	for *cpcB3A3*
FDGVPA	- X03101; X06085	for *gvpA1A2C*.

A

```
                          .          20         .          40         .
apcA1.AA  --MSIVTKSI VNADAEARYL SPGELDRIKS FVSGGERRLR IAQILTENRE
apcA2.AA  --MSIITKMI LNADAEVRYL TPGELDQINI FVKSSQRRLQ LVEALTQSRA
apcD.AA   --MTVISQVI LQADDELRYP SSGELKSIQA FMQTGVKRTR IASTLAENEK
cpcA1.AA  -MKTPLTEAV AAADSQGRFL SSTEIQTAFG RFRQASASLA AAKALTEKAS
cpcA2.AA  -MKTPLTEAV ATADSQGRFL SSTELQVAFG RFRQASASLD AAKALSSKAN
cpcA3.AA  MTKTPLTEAV VSADSQGRFL S-TELQVAFG RFRQAGSSLE AAKALSKKAS
cpeA.AA   -MKSVVTTVI AAADAAGRFP STSDLESVQG SIQRAAARLE AAEKLANNID

apcB1.AA  MAQDAITSVI NSADVQGKYL DSAALDKLKG YFGTGELRVR AASTISANAA
cpcB1.AA  -MLDAFAKVV SQADARGEYL SGSQIDALSA LVADGNKRMD VVNRITGNSS
cpcB2.AA  -MLDAFTKVV SQADTRGAYI SDAEIDALKT MVAAGSKRMD VVNRITGNAS
cpcB3.AA  MVQDAFSKVV SQADARGEYL SDGQLDALIN LVKEGNKRVD VVNRISSNAS
cpeB.AA   -MLDAFSRAV VSADASTSTV SD--IAALRA FVASGNRRLD AVNAIASNAS

                          .          70         .          90         .
apcA1.AA  RLVKQAGEQV FQKRPDVVSP GGNAYGQEL- -TATCLRDLD YYLRLVTYGI
apcA2.AA  TIVKQAGKDI FQRFPRLVAP GGNAYGENM- -TATCLRDMD YYLRLITYSV
apcD.AA   KIVQEATKQL WQKRPDFISP GGNAYGERQ- -RSLCIRDFG WYLRLITYGV
cpcA1.AA  SLASGAANAV YSKFPYTTSQ NGPNFASTQT GKDKCVRDIG YYLRMVTYCL
cpcA2.AA  SLAQGAVNAV YQKFPYTTQM QGKNFASDQR GKDKCARDIG YYIRIVTYCL
cpcA3.AA  SLAEAAANAV YQKFPYTTTT SGPNYASTQT GKDKCVRDIG YYLRIVTYGL
cpeA.AA   AVATEAYNAC IKKYPYLNNS GEANSTDTF- -KAKCARDIK HYLRLIQYSL

apcB1.AA  AIVKEAVAKS LL-YSDVTRP GGNMYTTRR- -YAACIRDLD YYFRYATYAM
cpcB1.AA  TIVANAARSL FAEQPGLIAP GGNAYTSRR- -MAACLRDME IILRYVTYAI
cpcB2.AA  TIVANAARAL FEEQPQLIAP GGNAYTNRR- -MAACLRDME IILRYVTYAV
cpcB3.AA  SIVANAARSL FAEQPQLIAP GGNAYTSRR- -AAACVRDLE IILRYVTYAI
cpeB.AA   CMVSDAVAGM ICENQGLIQA GGSCYPNRR- -MAACLRDAE IVLRYVTYAL

                          .         120         .         140         .
apcA1.AA  VSGDVTPIEE IGVIGAREMY KSLGTPIEGI TEGIRALKSG ASSLLSGED-
apcA2.AA  AAGDTTPIQE IGIVGVRQMY RSLGTPIDAV AESVRAMKNI TTSMLSGED-
apcD.AA   LAGDKEPIEK IGLIGVREMY NSLGVPVPGM VEAIASLKKA ALDLLSAED-
cpcA1.AA  VVGGTGPLDD YLIGGIAEIN RTFDLSPSWY VEALKYIKAN HG--LSGDP-
cpcA2.AA  VAGGTGPLDD YLIGGLAEIN RTFDLSPSWY VEALKYIKAN HG--LSGDP-
cpcA3.AA  VVGGTGPIDD YLIGGLAEIN RTFELSPSWY IEALKYIKAN HG--LSGDP-
cpeA.AA   VVGGTGPLDE WGIAGQREVY RALGLPTAPY VEALSFARNR GCA-PRDMSA

apcB1.AA  LAGDPSILDE RVLNGLKETY NSLGVPVSST VQAIQAIKEV TASLVGSDA-
cpcB1.AA  FAGDASVLDD RCLNGLKETY LAIGTPGSSV AVGVQKMKDA ALAIAGDTNG
cpcB2.AA  FAGDASVLDD RCLNGLRETY QALGVPGASV STGVQKMKEA AIAIANDPSG
cpcB3.AA  FAGDASVLDD RALNGLRETY LALGTPGASV AVGIQKLKES SIAIANDPNG
cpeB.AA   LAGDASVLDD RCLNGLKETY AALGVPTTST VRAVQIMKAQ AAAHIQDTPS

                         170
apcA1.AA  ---------- ----------- ---AAEAGSY FDYVVGALS
apcA2.AA  ---------- ----------- ---ASEVGTY FDYLITNLQ
apcD.AA   ---------- ----------- ---AAEASPY FDYIIQAMS
cpcA1.AA  ---------- ----------- ---AVEANSY IDYAINALS
cpcA2.AA  ---------- ----------- ---AVEANSY IDYAINALS
cpcA3.AA  ---------- ----------- ---AVEANSY IDYIINALS
cpeA.AA   QA-------- ----------- ---LTEYNAL LDYAINSLS

apcB1.AA  ---------- ----------- ---GKEMGVY LDYISSGLS
cpcB1.AA  ITRGD----- ---------C ASLMAEVASY FDKAASAVA
cpcB2.AA  VTRGD----- ---------C SSLMSELGSY FDRAAAAVG
cpcB3.AA  ITRGD----- ---------C SSLIAEVSGY FDRAAAAVA
cpeB.AA   EARAGAKLRK MGTPVVEDRC ASLVAEASSY FDRVISALS
```

B

```
                        .          20         .          40         .
apcC.AA  MARLFKVTAC VPSQTRIRTQ RELQNTYFTK LVPFENWFRE QQRIMKMGGK
         IVKVELATGK QGTNTGLL

                        .          20         .          40         .
cpcL.AA  MAPLTEASRL GVRPFADSDK VELRFVKTAE EVRSVIWSAY RQVLGNEHLF
cpcM.AA  MPITTAASRL GTSAFSNAAP IELRSNTNKA EIAQVIAAIY RQVLGNVIVL
cpcH.AA  MTSSTAARQL GFEPFASTAP TELRASS--- DVPAVIHAAY RQVFGNDHVM
cpcI.AA  MPITSAASRL GTTAYQ-TNP IELRPNWTAE DAKIVIQAVY RQVLGNDYLM
                        .          70         .          90         .
cpcL.AA  ESERLSSAES LLQQAQISVR DFVRAIAQSE LYRQKFFYSN SQVRFIELNY
cpcM.AA  QSERLKGLES LLTNGNITVQ EFVRQLAKSI YTSSFL
cpcH.AA  QSERLTSAES LLQQGNISVR DFVRLLAQSE LYRQKFFYST PQVRFIELNY
cpcI.AA  QSERLTSLES LLTNGKLSVR DFVRAVAKSE LYKTKFLYPH FQTRVIELNF
                        .         120         .         140         .
cpcL.AA  KHLLGRAPYD ESEIAYHVDI YTSQGYEAEI NSYIDSVEYQ QNFGDSIVPY
cpcM.AA
cpcH.AA  KHLNGRAPYD ESEISYHVNL YTEKGYEAEI NSYIDSAEYQ ESFGERIVPH
cpcI.AA  KHLLGRAPYD ESEVIEHLDR YQNQGFDADI DSYIDSAEYD TYFGDSIVPY
                        .         170         .         190         .
cpcL.AA  YRGYQTT-VG QKTAGFPRFF QLYRGYANRD R-QNKSKGQ- LTWDLAKNLV
cpcM.AA
cpcH.AA  YRGFETQ-PG QKTVGFNRMF QIYRGYANSD RSQGKNKSAW LTQDLALNLA
cpcI.AA  YRDLVTTGVG QRTVGFTRMF RLYRGYANSD RSQLAGSSSR LASDLATNSA
                        .         220         .         240         .
cpcL.AA  SPIY-PADA- ---------- -----GSLTG VSTGNRGNTY RIRTTQAASP
cpcM.AA
cpcH.AA  SNIQTPNFG- ---------- -----KGLTG VVAGDRGQLY RVRVIQADRG
cpcI.AA  TAIIAPSGGT QGWSYLPSKQ GTAPSRTFGR SSQGSTPRLY RIEVTGISLP
                        .         270         .
cpcL.AA  NSPRIRQSIS EVVVPFDQLS NLLQQLNRQG GKVISIALS
cpcM.AA
cpcH.AA  RTTQIRRSIQ EYLVSYDQLS PTLQRLNQRG SRVVNISPA
cpcI.AA  RYPKVRRSNK EFIVPYEQLS STLQQINKLG GKVASITFAQ

                        .          20         .          40         .
cpcD.AA  MLGSVLTRRS SSGSDNRVFV YEVEGLRQNE QTDNNRYQIR NSSTIEIQVP
         YSRMNEEDRR ITRLGGRIVN IRPAGENPTE DASEN
```

C

```
                          20                  40
cpcE.AA  MYRHLSEGIE DHREQEQKVE NAANIQDDNQ LTVEQAIANL QGEDLGLRVY
         AAWWLGRFRV DAPEAIDVLI QALEDEDDRT NVGGYPLRRN AARALGKLGE
         KRAVPALIKA LECSDFYVRE AAAQSLEMLG DSSSIPRLIE LLNDQVPGTL
         PAPEPPQLTQ PFDAIIEALG TLGASDAIPI IQEFLEHTVP RIQYAAARAM
         YQLTSESTAG YNQYGDRLVQ ALAQDDLQLR RAVLSDLGAI GYLPAAEAIP
         DTLAENSLKL ISLKGLLEKQ FQPTKPEDLS PGAIKVMQLM DALL

                          20                  40
cpcF.AA  MPDSLNSLIR AVEDANSSIL LQEAVKNLAA ARLEGAIPTL IAALSYNNPG
         AAVAAVDGLI QIGEPAVPSL LELLDMHNYT ARSWAIRTLA GIGDPRGLVT
         LLGAATADFA LSVRRAAAKG LGMMKWHWFP EELLEIAQAE AMEALLFVAQ
         EDEEWVVRYS AIVGLQFLAN AIAVSHPDWR SQILSNFEQI AAKEESWPVR
         ARVLLAQQEL QQITATIPTQ DIENRPSPLS SMDWQKIMED LYGRKGQERL
         VFAEGDPRRI

                          20                  40
orfY.AA  MDKRFFNFFN LTEDQAIALL DTPQDQLSEN DSRYIAASHL VNFPTERSIN
         ALIRAVQQTD PSLDNRIVRR KSVETLGRLK ATTALPFIRI CLFDEDCYTV
         ENAAWAIGEI GTQDTDILED VAQLLEKPGQ TYRVIIHTLT KPNYQPALER
         IRKFVNDSDP PTASAAIAAV CRLTGDYSQM AKVVQILQHP NVLGRRLSIQ
         DLMDARYYDA IPDIAKCPVS LVFRLRGLRT LAEAGISEGA ITFAKIQPYL
         EQTLYDHPQD LNLVHSYDRL PTLEILIRGL YETDFGRCYL ATKTILEHYA
         DAAAEALFAT YAAEANNDYG AHFHVIKLFG WLKHAPAYDL IVEGLHNKQP
         QFQKSRAAAA IALAELGDPK AIPELKACLE TKIWDLKYAT LMALEKLGDI
         SEHKQAAQDS DWLIARKASS TLKNQEITA

                          20                  40
orfZ.AA  MPTTEELFQQ LKHPNPHLRD QAMWELAENP DETTIPRLMS ILDEEDTTYR
         RAAVKALGAI GPDAITPLVQ ALLNSDNVTV RGSAAKALAQ VAINHPDVPF
         AAEGVQGLKT ALDDPNPVVH IAAVMALGEI GSPVVDVLIE ALQTTDNPAL
         GISIVNALGS IGDSRGVEVL QSLIENESTD SYVRESATSA LSRLEMTTKF
         QRGEK
```

Fig. 11. Predicted amino acid sequences of the polypeptides involved in the formation of phycobilisomes in *Calothrix* 7601. The single-letter code is that recommended by IUPAC/IUB. Some of the previously published gene designations have been changed according to the nomenclature rules proposed for cyanobacterial genes (Houmard and Tandeau de Marsac 1988).

A. Phycobiliprotein subunits: αAP1 (*apcA1*.AA), αAP2 (*apcA2*.AA), αAPB (*apcD*.AA), αPC1 (*cpcA1*.AA), αPC2 (*cpcA2*.AA), αPC3 (*cpcA3*.AA), αPE (*cpeA*), βAP1 (*apcB1*.AA), βPC1 (*cpcB1*.AA), βPC2 (*cpcB2*.AA), βPC3 (*cpcB3*.AA) and βPE (*cpeB*).

B. Linker polypeptides: $L_C^{7.8}$ (*apcC*.AA), L_R^{38}-like polypeptide (*cpcL*.AA), L_R-like polypeptide (*cpcM*.AA), L_R^{38} (*cpcH*.AA), L_R^{39} (*cpcI*.AA) and $L_R^{9.7}$ (*cpcD*.AA). The last three sequences are taken from Lomax et al. 1987.

C. Unassigned open reading frames.

```
                          20                  40
gvpA1.AA MAVEKTNSSS SLAEVIDRIL DKGIVVDAWV RVSLVGIELL AIEARIVIAS
         VETYLKYAEA VGLTQSAAVP A

                          20                  40
gvpC.AA  MTPLMIRIRQ EHRGIAEEVT QLFKDTQEFL SVTTAQRQAQ AKEQAENLHQ
         FHKDLEKDTE EFLTDTAKER MAKAKQQAED LFQFHKEMAE NTQEFLSETA
         KERMAQAQEQ ARQLREFHQN LEQTTNEFLA DTAKERMAQA QEQKQQLHQF
         RQDLFASIFG TF
```

Fig. 12. Predicted amino acid sequences of the polypeptides involved in gas vesicle formation in *Calothrix* 7601. The amino acid sequence of the *gvpA2* gene is strictly identical to that of *gvpA1*. The single letter code is that recommended by IUPAC/IUB.

References

Allen JF, Sanders CE and Holmes NG (1985) Correlation of membrane protein phosphorylation with excitation energy distribution in the cyanobacterium *Synechococcus* 6301. FEBS Lett 193: 271–275

Anderson LK, Rayner MC, Sweet RM and Eiserling FA (1983) Regulation of *Nostoc* sp. phycobilisome structure by light and temperature. J Bacteriol 155: 1407–1416

Anderson LK and Eiserling FA (1986) Asymmetrical core structure in phycobilisomes of the cyanobacterium *Synechocystis* 6701. J Mol Biol 191: 441–451

Armstrong RE, Hayes PK and Walsby AE (1983) Gas vacuole formation in hormogonia of *Nostoc muscorum*. J Gen Microbiol 128: 263–270

Belasco JG, Beatty JT, Adams CW, von Gabain A and Cohen SN (1985) Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic *rxcA* transcript. Cell 40: 171–181

Belknap WR and Haselkorn R (1987) Cloning and light regulation of expression of the phycocyanin operon of the cyanobacterium *Anabaena*. EMBO J 6: 871–884

Bennett A and Bogorad L (1973) Complementary chromatic adaptation in a filamentous blue-green alga. J Cell Biol. 58: 419–435

Brawerman G (1987) Determinants of messenger RNA stability. Cell 48: 5–6

Bryant DA, de Lorimier R, Guglielmi G, Stirewalt VL, Dubbs JM, Illman B, Gasparich G, Buzby JS, Cantrell A, Murphy RC, Gingrich J, Porter RD and Stevens SE Jr (1986) The cyanobacterial photosynthetic apparatus: a molecular genetic analysis. In: Youvan DC and Daldal F (eds) Microbial Energy Transduction. Genetics, Structure, and Function of Membrane Proteins. pp 39–46. Cold Spring Harbor: Cold Spring Harbor Laboratory

Bryant DA (1988) Phycobilisomes of *Synechococcus* sp. PCC 7002, *Pseudanabaena* sp. PCC 7409, and *Cyanophora paradoxa*: an analysis by molecular genetics. In: Scheer H and Schneider S (eds) Photosynthetic Light-harvesting Systems. Structure and function. Berlin: de Gruyter & Co, in press

Capuano V, Mazel D, Tandeau de Marsac N and Houmard J (1988) Complete nucleotide sequence of the red-light specific set of phycocyanin genes from the cyanobacterium PCC 7601. Nucleic Acids Res 16: 1626

Cech TR, Tanner NK, Tinoco I Jr, Weir BR, Zuker M and Perlman PS (1983) Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: Structural homology with fungal mitochondiral intervening sequences. Proc Natl Acad Sci USA 80: 3903–3907

Cohen-Bazire G and Bryant DA (1982) Phycobilisomes: composition and structure. In: Carr NG and Whitton BA (eds) The Biology of Cyanobacteria, pp 143–190. Oxford: Blackwell Scientific Publications

Cone JW and Kendrick RE (1986) Photocontrol of seed germination. In: Kendrick RE and Kronenberg GHM (eds) Photomorphogenesis in plants, pp 443–465. Dordrecht: Martinus Nijhoff Publishers

Conley PB, Lemaux PG and Grossman AR (1985)· Cyanobacterial light-harvesting complex subunits encoded in two red light-induced transcripts. Science 230: 550–553

Conley PB, Lemaux PG, Lomax TL and Grossman AR (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

Csiszàr K, Houmard J, Damerval T and Tandeau de Marsac N (1987) Transcriptional analysis of the cyanobacterial *gvpABC* operon in differentiated cells: occurrence of an antisense RNA complementary to three overlapping transcripts. Gene 60: 29–38

Damerval T, Houmard J, Guglielmi G, Csiszàr K and Tandeau de Marsac N (1987) A developmentally regulated *gvpABC* operon is involved in the formation of gas vesicles in the cyanobacterium *Calothrix* 7601. Gene 54: 83–92

DasSarma S, Damerval T, Jones JG and Tandeau de Marsac N (1987) A plasmid-encoded gas vesicle protein gene in a halophilic archaebacterium. Mol Microbiol 1: 365–370

De Lorimier R, Bryant DA, Porter RD, Liu W-Y, Jay E and Stevens SE Jr (1984) Genes for the α and β subunits of phycocyanin. Proc Natl Acad Sci USA 81: 7946–7950

Elmorjani K and Herdman M (1987) Metabolic control of phycocyanin degradation in the cyanobacterium *Synechocystis* PCC 6803: a glucose effect. J Gen Microbiol 133: 1685–1694

Flores E and Wolk CP (1985) Identification of facultatively heterotrophic, N_2-fixing cyanobacteria able to receive plasmid vectors from *Escherichia coli* by conjugation. J Bacteriol 162: 1339–1341

Foulds IJ and Carr NG (1977) A proteolytic enzyme degrading phycocyanin in the cyanobacterium *Anabaena cylindrica*. FEMS Microbiol Lett 2: 117–119

Füglistaller P, Suter F and Zuber H (1983) The complete amino-acid sequence of both subunits of phycoerythrocyanin from the thermophilic cyanobacterium *Mastigocladus laminosus*. Hoppe-Seyler's Z Physiol Chem 364: 691–712

Füglistaller P, Rümbeli R, Suter F and Zuber H (1984) Minor polypeptides from the phycobilisome of the cyanobacterium *Mastigocladus laminosus*. Isolation, characterization and amino-acid sequences of a colourless 8.9 kDa polypeptide and of a 16.2 kDa phycobiliprotein. Hoppe-Seyler's Z Physiol Chem 365: 1085–1096

Füglistaller P, Suter F and Zuber H (1985) Linker polypeptides of the phycobilisome from the cyanobacterium *Mastigocladus laminosus*: amino-acid sequences and relationships. Biol Chem Hope-Seyler 366: 993–1001

Füglistaller P, Suter F and Zuber H (1986) Linker polypeptides of the phycobilisome from the cyanobacterium *Mastigocladus laminosus*. II Amino-acid sequences and functions. Biol Chem Hoppe-Seyler 367: 615–626

Gendel S, Ohad I and Bogorad L (1979) Control of phycoerythrin synthesis during chromatic adaptation. Plant Physiol 64: 786–790

Gingrich JC, Lundell DJ and Glazer AN (1983) Core substructure in cyanobacterial phycobilisomes. J Cell Biochem 22; 1–14

Glazer AN (1982) Phycobilisomes: structure and dynamics. Ann Rev Microbiol 36: 173–198

Glazer AN (1984) Phycobilisome. A macromolecular complex optimized for light energy transfer. Biochim Biophys Acta 768: 29–51

Glazer AN (1985) Light harvesting by phycobilisomes. Ann Rev Biophys Biophys Chem 14: 4–77

Glazer AN and Clark JH (1986) Phycobilisomes. Macromolecular structure and energy flow dynamics. Biophys J 49: 115–116

Glazer AN (1987) Phycobilisomes: assembly and attachment. In: Fay P and Van Baalen C (eds) The Cyanobacteria, pp 69–94. Amsterdam: Elsevier Science Publishers

Glick RE and Zilinskas BA (1982) Role of the colorless polypeptides in phycobilisome reconstitution from separated phycobiliproteins. Plant Physiol 69: 991–997

Grossman AR, Lemaux PG and Conley PB (1986) Regulated synthesis of phycobilisome components. Photochem Photobiol 44: 827–837

Hayes PK, Walsby AE and Walker JE (1986) Complete amino acid sequence of cyanobacterial gas vesicle protein indicates a 70-residue molecule that corresponds in size to the crystallographic unit cell. Biochem J 236: 31–-36

Houmard J, Mazel D, Moguet C, Bryant DA and Tandeau de Marsac N (1986) Organization and nucleotide sequence of genes encoding core components of the phycobilisomes from *Synechococcus* 6301. Mol Gen Genet 205: 404–410

Houmard J, Capuano V, Coursin T and Tandeau de Marsac N (1988) Isolation and molecular characterization of the gene encoding allophycocyanin B, a terminal energy acceptor in cyanobacterial phycobilisomes. Mol Microbiol 2: 101–107

Houmard J and Tandeau de Marsac N (1988) Cyanobacterial Genetic Tools: Current Status. In: Packer L and Glazer AN (eds) Methods in Enzymology: Cyanobacteria. Academic Press, (in press)

Jost M and Jones DD (1970) Morphological parameters and macromolecular organization of gas vacuole membranes of *Microcystis aeruginosa* Kuetz. emend. Elenkin. Can J Microbiol 16: 159–164

Klotz AV, Leary JA and Glazer AN (1986) Post-translational methylation of asparaginyl residues. Identification of β-71γ-N-methylasparagine in allophycocyanin. J Biol Chem 261: 15891–15894

Lau RH, MacKenzie MM and Doolittle WF (1977) Phycocyanin synthesis and degradation in the blue-green bacterium *Anacystis nidulans*. J Bacteriol 132: 771–778

Lomax TL, Conley PB, Schilling J and Grossman AR (1987) Isolation and characterization of light-regulated phycobilisome linker polypeptide genes and their transcription as a polycistronic mRNA. J Bacteriol 169: 2675–2684

Mazel D, Guglielmi G, Houmard J, Sidler W, Bryant DA and Tandeau de Marsac N (1986) Green light induces transcription of the phycoerythrin operon in the cyanobacterium *Calothrix* 7601. Nucleic Acids Res 14: 8279–8290

Mazel D, Houmard J and Tandeau de Marsac N (1988) A multigene family in *Calothrix* sp. PCC 7601 encodes phycocyanin, the major component of the cyanobacterial light harvesting antenna. Mol Gen Genet 211: 296–304

Minami Y, Yamada F, Hase T, Matsubara H, Murakami A, Fujita Y, Takao T and Shimonishi Y (1985) Amino acid sequences of allophycocyanin α- and β-subunits isolated from *Anabaena cylindrica* FEBS Lett 191: 216–220

Newbury SF, Smith NH, Robinson EC, Hiles ID and Higgins CF (1987) Stabilization of translationally active mRNA by prokaryotic REP sequences. Cell 48: 297–310

Ohad I, Clayton RK and Bogorad L (1979) Photoreversible absorbance changes in solutions of allophycocyanin purified from *Fremyella diplosiphon*: temperature dependence and quantum efficiency. Proc Natl Acad Sci USA 76: 5655–5659

Ohad I, Schneider HAW, Gendel S and Bogorad L (1980) Light-induced changes in allophycocyanin. Plant Physiol 65: 6–12

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

Platt T (1986) Transcription termination and the regulation of gene expression. Ann Rev Biochem 55: 339–372

Raps S, Kycia JH, Ledbetter MC and Siegelman HW (1985) Light intensity adaptation and phycobilisome composition of *Microcystis aeruginosa*. Plant Physiol 79: 983–987

Riethman HC, Mawhinney TP and Sherman LA (1987) Phycobilisome-associated glycoproteins in the cyanobacterium *Anacystis nidulans* R2. FEBS Lett 215: 209–214

Rippka R and Herdman M (1985) Division patterns and cellular differentiation in cyanobacteria. Ann Microbiol (Inst Pasteur) 136A: 33–39

Rümbeli R, Suter F, Wirth M, Sidler W and Zuber H (1987) γ-*N*-Methylasparagine in phycobiliproteins from the cyanobacteria *Mastigocladus laminosus* and *Calothrix*. FEBS Lett 221: 1–2

Sidler W, Kumpf B, Rüdiger W and Zuber H (1986) The complete amino-acid sequence of C-phycoerythrin from the cyanobacterium *Fremyella diplosiphon*. Biol Chem Hoppe-Seyler 367: 627–642

Stanier RY and Cohen-Bazire G (1977) Phototrophic prokaryotes: The Cyanobacteria. Ann Rev Microbiol 31: 225–274

Stark GR and Wahl GM (1984) Gene amplification. Ann Rev Biochem 53: 447–491

Stevens SE Jr, Balkwill DL and Paone DAM (1981) The effects of nitrogen limitation on the ultrastructure of the cyanobacterium *Agmenellum quadruplicatum*. Arch Microbiol 130: 204–212

Suter F, Füglistaller P, Lundell DJ, Glazer AN and Zuber H (1987) Amino acid sequences of α-allophycocyanin B from *Synechococcus* 6301 and *Mastigocladus laminosus*. FEBS Lett 217: 279–282

Tandeau de Marsac N (1977) Occurrence and nature of chromatic adaptation in cyanobacteria. J Bacteriol 130: 82–91

Tandeau de Marsac N (1983) Phycobilisomes and complementary chromatic adaptation in cyanobacteria. Bull Inst Pasteur 81: 201–254

Tandeau de Marsac N, Mazel D, Bryant DA and Houmard J (1985) Molecular cloning and nucleotide sequence of a developmentally regulated gene from the cyanobacterium *Calothrix* PCC 7601: a gas vesicle protein gene. Nucleic Acids Res 13: 7223–7236

Walker JE, Hayes PK and Walsby AE (1984) Homology of gas vesicle proteins in cyanobacteria and halobacteria. J Gen Microbiol 130: 2709–2715

Walsby AE (1978) The gas vesicles of aquatic prokaryotes. Symp Soc Gen Microbiol 28: 327–358

Walsby AE (1981) Cyanobacteria: planktonic gas-vacuolate forms. In: Starr M, Stolp H, Trüper H, Balows A and Schlegel HG (eds) The Prokaryotes, pp 224–235. New York: Springer-Verlag

Walsby AE (1987) Mechanisms of buoyancy regulation by planktonic cyanobacteria with gas vesicles. In: Fay P and Van Baalen C (eds) The Cyanobacteria, pp 377–392. Amsterdam: Elsevier Science Publishers

Wolk CP, Vonshak A, Kehoe P and Elhai J (1984) Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria. Proc Natl Acad Sci USA 81: 1561–1565

Wong HC and Chang S (1986) Identification of a positive retroregulator that stabilizes mRNAs in bacteria. Proc Natl Acad Sci USA 83: 3233–3237

Wood NB and Haselkorn R (1980) Control of phycobiliprotein proteolysis and heterocyst differentiation in *Anabaena*. J Bacteriol 141: 1375–1385

Yamanaka G and Glazer AN (1980) Dynamic aspects of phycobilisome structure. Phycobilisome turnover during nitrogen starvation in *Synechococcus* sp. Arch Microbiol 124: 39–47

Yamanaka G and Glazer AN (1981) Dynamic aspects of phycobilisome structure: modulation of phycocyanin content of *Synechococcus* phycobilisomes. Arch Microbiol 130: 23–30

Zilinskas BA, Chen KH and Howell DA (1987) Cloning genes for phycobilisome core components. Plant Physiol (Suppl) 83: 60