Gene transcription in filamentous cyanobacteria

Jean Houmard

Tel: +33 1 45 68 84 16. Fax: +33 1 40 61 30 42. e-mail jhoumard@pasteur.fr

Département de BGM, CNRS, URA 1129 Physiologie Microbiennne, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cedex 15, France

Keywords: Filamentous cyanobacteria, transcription, RNA polymerase, DNA-binding protein, signal transduction

General features of cyanobacteria

Cyanobacteria are probably the most diverse group of prokaryotes in the number of species, type of habitats, morphology and physiological properties. Their common feature is O₂-evolving photosynthesis, performed by mechanisms similar to those of algae and higher plants. Because cyanobacteria can live in many very different ecological niches, they have experienced different environmental stresses and may, thus, have evolved specialized adaptation mechanisms. Some strains are strict photoautotrophs, others can alternatively use exogenous carbon sources such as glucose. A rather large number of cyanobacteria may reduce atmospheric dinitrogen, either in specialized cells for those strains that differentiate heterocysts, or in vegetative cells of nonheterocystous strains. Akinete, hormogonium and baeocyte formation have been described for some strains (for a review, see Tandeau de Marsac & Houmard, 1993).

Filamentous cyanobacteria form three of the five different provisional taxonomic sections (III to V) defined by Rippka et al. (1979): Section III groups the strains which do not differentiate heterocysts; Section IV comprises most of the strains that can form heterocysts, while Section V accommodates the heterocystous strains which in addition have true branching filaments. Almost all the studies dealing with gene transcription in filamentous cyanobacteria have been performed with a very limited number of strains, mostly of section IV, such as Anabaena Nostoc PCC 7120 and PCC 7937, Nostoc commune UTEX 584 and Calothrix PCC 7601.

I shall not present an exhaustive review on transcription in cyanobacteria in general, since the topic was recently covered by Curtis & Martin (1994). I shall restrict the discussion to our present knowledge of filamentous cyanobacteria, but it is worth noting that no major differences have yet been found in the mechanisms regulating transcription in unicellular and filamentous strains. Regarding housekeeping functions such as DNA replication, translation, nucleotide and amino acid biosyntheses, very little has been defined. Most investigators have focused on the properties that distinguish cyanobacteria from the other prokaryotes, implicitly assuming that the basic phenomena described for Escherichia coli and/or Bacillus subtilis will apply to the cyanobacteria. However, the diversity and the specificity of cyanobacteria have to be kept in mind, especially when considering regulatory mechanisms. After a brief description of the cyanobacterial transcription machinery and of the genes that encode it, I shall review the signals that have been recognized as important in the control of gene transcription, and finally I shall present examples of transcriptional regulation found in filamentous cyanobacteria.

The transcription machinery

The first report on the purification and characterization of a DNA-dependent RNA polymerase from a cyanobacterium was from the unicellular strain Anacystis nidulans, now designated Synechococcus sp. (Herzfeld & Zillig, 1971). Since then the enzyme has been characterized from two filamentous strains, Anabaena Nostoc PCC 7120 (Schneider et al., 1987) and Calothrix PCC 7601, also called Fremyella diplosiphon UTEX 481 (Miller & Bogorad, 1978; G. Schyns & J. Houmard, unpublished data). Their polypeptide compositions resemble that of the well-characterized prokaryotic enzymes, with β, β′, α and σ subunits. For the E. coli-type RNA polymerase, the main features are: (i) individual subunits are inactive; (ii) β is the catalytic subunit for RNA synthesis; (iii) β′ is a basic protein able to non-specifically bind to DNA; (iv) two α subunits, together with a β subunit, give an αβ complex which in turn binds β′ to form the core enzyme (E); and (v) σ subunits confer promoter specificity to the holoenzyme (Eσ) (Ishihama, 1992). The cyanobacterial RNA polymerases very likely share these properties but they have an additional subunit, γ, which has been found in all the cyanobacteria examined to date (Schneider & Hasel-
The stoichiometry for E is $\alpha_2\beta\beta'\gamma$ and $\alpha_2\beta\beta''\gamma\epsilon$ for Ea. The additional $\gamma$ subunit is equivalent to the $\beta'$ polypeptide that has been described for DNA-dependent RNA polymerases of chloroplasts which, as known at present, have a composition that very much resembles the cyanobacterial enzyme (Bogorad, 1991; Igloi & Kossel, 1992). In maize chloroplasts, it comprises polypeptides of 180 kDa ($\beta'$), 120 kDa ($\beta$), 78 kDa ($\beta' = \text{cyanobacterial} \gamma$) and 38 kDa ($\alpha$). Polypeptides, designated SLF (sigma-like factors), have also been described but none of them have been shown unequivocally to be the functional equivalents of bacterial $\sigma$ proteins (Bogorad, 1991; Tiller et al., 1991). Both the cyanobacterial $\gamma$ and the chloroplast $\beta'$ subunits are homologous to the N-terminal part of the $\beta'$ subunit of the other eubacteria examined so far. The N-terminal sequence of what is called $\beta'$ in cyanobacteria and $\beta''$ in chloroplasts is homologous to the C-terminal half of the ‘classical’ $\beta'$ of eubacterial RNA polymerase. Thus, both the cyanobacterial $\beta'$ and the plastidial $\beta''$ possess an additional C-terminal extension of more than 50 kDa. Such a splitting of RNA polymerase components has also been found in archaeobacteria where the homology to the E. coli $\beta'$ subunit is divided between the $\alpha'$ and $\alpha''$ (also designated A and C) components of the RNA polymerase (Zillig et al., 1992). The analogy between these polypeptides is reinforced by antibodies raised against the cyano- bacterial $\gamma$ subunit which cross-react with a chloroplast $\beta'$ polypeptide made in E. coli and with archaeobacterial $\Lambda'$ polypeptides (Bergsland & Haselkorn, 1991). Many archaeobacteria also differ from the other bacteria by having the $\beta$ subunit equivalent also split, leading to $B'$ and $B''$ subunits (Zillig et al., 1992).

In cyanobacteria, the genes encoding the subunits of E are designated $rpo$, $sig$ being used for the genes that code for $\sigma$ polypeptides. The $rpoBC1C2$ cluster (genes encoding $\beta$, $\gamma$ and $\beta'$, respectively) has been isolated from N. commune UTEX 584 (Xie et al., 1989) and Anabaena/Nostoc PCC 7120 (Bergsland & Haselkorn, 1991). Although not yet characterized, the $rpoA$ gene encoding the $\alpha$ subunit is not adjacent to this cluster. Three $sig$ genes, $sigA$, $sigB$ and $sigC$, have already been characterized from Anabaena/Nostoc PCC 7120. The $sigA$ gene most probably encodes the major $\sigma$ factor. The function of the other two gene products remains to be determined since they can be inactivated without affecting cell viability, heterocyst differentiation or $N_2$ fixation under laboratory conditions (Brahamsha & Haselkorn, 1991, 1992).

As mentioned above, the cyanobacterial and chloroplast RNA polymerases resemble each other, but other features concerning cyano bacteria and chloroplasts can be mentioned. In contrast to nuclear genes, chloroplast genes are often part of polycistronic transcription units similar to those found in bacteria (Weil, 1987). Reciprocally, introns have been recognized for some years now in chloroplast genes, but some cyanobacterial genes also possess intervening sequences. In addition to the rare group I intron described for a few tRNA genes, two group II self-splicing introns have been found in Calothrix PCC 7601 (Ferat & Michel, 1993). Finally, it is worth noting that control of transcription by reversible phosphorylation has now also been demonstrated in chloroplasts (Tiller & Link, 1993). Altogether, these data reinforce the endosymbiotic theory on the origin of chloroplasts. A major difference between chloroplasts and cyanobacteria resides in the presence of a phage-type DNA-dependent RNA polymerase which seems to coexist with the bacterial type enzyme, at least in the plastids of young leaves of spinach (Lerbs-Mache, 1993). However, this enzyme may well be nuclear-encoded if it corresponds to the RNA polymerase activity found in cells that contain undifferentiated plastids (Hess et al., 1993). In agreement with the similarity in the transcription apparatus of plastids and cyanobacteria, it has been shown that the promoter $P_{paA}$, derived from the chloroplast of Amaranthus hybridus, efficiently drives the expression of an antibiotic cassette cloned downstream of it, in Anabaena/Nostoc PCC 7120 (Elhai & Wolk, 1988).

Transcription signals

Transcription initiation

It is now well-established that $\sigma$ factors play a major role in prokaryotic cells by conferring specificity to the RNA polymerase. The $\sigma$ factors recognize a DNA sequence as being a promoter and help the RNA polymerase to properly initiate transcription. In some instances, it has been possible to assign a consensus recognition sequence to a given $\sigma$ factor, TGTACG(N$_1$)$_2$TATAAT (centred at about $-35$ and $-10$) for E. coli $\sigma^32$ and GG(N$_{19}$)GC (at $-24$ and $-12$) for $\sigma^{34}$. At present, no such consensus sequences have been determined for cyanobacterial promoters, in the most part because only a few transcription start sites have been mapped and thus the size of the sample is too small to generate statistically significant data. Moreover, expression of many of the genes that have been studied is influenced by very different environmental parameters and may involve unique signal-response mechanisms. In any case, the concept of a strict recognition of a given DNA sequence by a specific $\sigma$ factor must be reexamined in view of the recent results of Tanaka et al. (1993). These authors have shown that both $\sigma^70$ and $\sigma^2$ holoenzymes purified from E. coli may properly initiate transcription at promoters which, on a strict basis of sequence similarities, would have been specifically assigned to the $\sigma^70$ promoter family. Conversely, the assignment to this family does not mean constitutive expression, since many bacterial genes that possess such promoters are regulated via operator regions that are the targets for repressors or activators such as CRP or FNR (Kolb et al., 1993).

The RNA polymerase purified from Anabaena/Nostoc PCC 7120 cells grown under nitrogen-replete conditions has been used in in vitro transcription assays. For six out of seven genes examined, the transcription start points (tsp) determined for in vivo mRNAs by S1 mapping and/or primer extension experiments have been reproduced in vitro (Schneider et al., 1991). For the glmA promoter region, three of the four tsp previously recognized in vegetative cells were functional in vitro, but the major tsp that was detected in cells grown under nitrogen-fixing conditions is not recognized by that purified form of
Operons

RNA polymerase, nor were the nif genes of this strain. Although the authors have stressed that the cyanobacterial polymerase prefers promoters for which the sequence is close to that of the canonical E. coli -35 and -10 of Esa70, they also observed that transcription at the psbA1 and rbcL5 promoters is enhanced by a component present in the cyanobacterial crude extract but not in the purified enzyme. Transcription appeared to be better in the E. coli system only for the hapB promoter (Schneider et al., 1991).

DNaseI protection experiments performed with the Calothrix PCC 7601 RNA polymerase revealed that the cyanobacterial enzyme behaves like the homologous bacterial DNA-dependent RNA polymerases. As with the E. coli and B. subtilis enzymes, the protection extends from about +15 to −40 (Sobczyk et al., 1993). Similar patterns have also been seen with a chloroplast enzyme (Zaitlin et al., 1989). The interactions between the promoter region and the RNA polymerase are thus expected not to vary greatly amongst organisms, although the role of the extra sequence of about 70 kDa present in the cyanobacterial beta and chloroplast beta prime subunits and its effect on the topology of the transcription apparatus [see the model proposed by Ishihama (1992) for the E. coli enzyme] remain to be clarified.

Transcription termination

Few 3' ends of transcripts have been mapped so far. The purified RNA polymerases of Anabaena/Nostoc PCC 7120 and Calothrix PCC 7601 recognize, in vitro, the transcription terminator of the coliphage T7 (Schneider & Haselkorn, 1991; G. Schyns & J. Houmard, unpublished data). In vivo, in almost every instance examined, the 3' extremities of the transcripts map a few nucleotides after possible stem and loop structures. However, per se, such putative structures do not seem to always lead to transcription termination. It has been observed that hairpins of high thermodynamic stability may often form in the intercistronic regions of polycistronic transcription units but they do not always prevent readthrough (Fig. 1). Transcripts of various sizes have often been found in Calothrix PCC 7601 that correspond either to part of or to a complete operon. Another example in that strain is the antisense RNA that is read through the structure present between gvpA1 and gvpA2. Finally, it must be stressed
that, in filamentous cyanobacteria, stretches of U like those often found in the so-called Rho-independent terminators of *E. coli* have seldom been found following the hairpins. Specific factors very likely operate in transcription termination in cyanobacteria, and functional homologues of Rho, for example, will certainly be discovered. The function of most of the possible stem and loop structures found downstream of the coding regions is probably more to act as barriers against 3′-5′ exonuclease activities than to be signals for transcription termination, *sensu stricto*.

**Fate of the transcripts**

Once made, the transcripts may either be directly translated or accumulate. The overall mRNA levels can be considered as the sum of two processes: (i) initiation frequency and elongation rate which define the rate of synthesis, that itself may depend on the presence of transcriptional activators or repressors; (ii) mRNA stability influenced by direct inactivation through endonucleolytic cleavage with enzymes equivalent to the RNaseIII, RNaseE, RNaseK and polynucleotide phosphorylase of *E. coli*, for example, and/or inactivation by a processive degradation from one end, and/or translation rate (for a review, see Petersen, 1992). Each of these mechanisms could be a target for controls.

The occurrence and role of nucleases in cyanobacteria is an undeveloped topic. A sugar-nonspecific enzyme, NucA, has recently been described in *Anabaena/Nostoc* PCC 7120 (Muro-Pastor et al., 1992). However, it is not possible at present to assign it any specific regulatory function since, at least *in vitro*, it may act as an endonuclease that degrades both linear and covalently closed circular plasmid DNA, as well as RNA. The RNA subunit of RNase P has also been characterized from *Anabaena/Nostoc* PCC 7120 and *Calothrix PCC 7601*, but it probably plays no role in transcript stability since the known physiological role of this endonuclease is to generate mature 5′-end of tRNAs from precursors of tRNAs (Vioque, 1992).

In contrast to *E. coli* or *B. subtilis*, most cyanobacterial genes characterized possess long transcribed but untranslated regions which do not appear to encode any polypeptide. For example, a length of more than 500 nt has been reported for the *sigA* and *rhlL* genes of *Anabaena/Nostoc* PCC 7120 (Brahamsha & Haselkorn, 1992; Curtis & Martin, 1994; Schneider et al., 1991), while values ranging from 23 to 362 nt have been found in front of the genes that encode the phycobiliproteins in *Calothrix PCC 7601* (Tandeau de Marsac et al., 1988). Roles in the stabilization of the mRNAs or in translational controls have been hypothesized for these sequences. Using wild-type or modified *psbB* promoter sequences in transcriptional fusions to the *cat* gene, Lang & Haselkorn (1991) indeed found regulatory elements in the DNA sequence downstream from the transcription start site of *psbB* in *Anabaena/Nostoc* PCC 7120. Interestingly, these elements have been shown to operate only *in vivo*, since the deletion of 192 bp downstream from the +1 has no effect on the transcripts generated *in vitro* but decreases the level of the *cat* gene product *in vivo* eightfold.

**Transcriptional regulation**

The occurrence of transcriptional regulations in filamentous cyanobacteria was first postulated based on the differences observed in gene expression upon addition of metabolic inhibitors such as rifampicin and chloramphenicol, antibiotics known to prevent RNA and protein synthesis, respectively (Gendel et al., 1979). Using gene isolation and sequencing techniques, specific hybridization probes have been designed which provide evidence at a molecular level that controls of RNA synthesis operate in the diverse processes mentioned below. Separation of total RNA by gel electrophoresis, transfer and immobilization of the RNAs to solid supports, followed by hybridization with specific probes, is the method of choice to evaluate the role of transcription in the regulation of the expression of a given gene. RNA blots not only allow estimation of the amount and of the physical integrity of the various mRNAs, but also of the relative ratios between transcripts that, although encoding the same gene product, may have different extremities. In parallel with RNA blot hybridizations, precise mappings of the 5′ and 3′ extremities of RNAs extracted from cells grown under well-defined conditions permit the location of the promoter and terminator regions on the DNA sequences.

For cyanobacteria, photon availability (often referred to as light intensity) is certainly one of the most important parameters of the environment. The role of the available photosynthetic photon flux density (PPFD) is quite pleiotropic and the molecular mechanism(s) by which it is sensed, as well as the pathway(s) by which the signal is transduced and reaches the transcription apparatus, remain to be clarified. The importance of the PPFD in regulating the expression of many of the genes related to photosynthesis has been recently reviewed (Tandeau de Marsac & Houmard, 1993) and thus I have selected only some of the examples of transcriptional regulation that have been recognized in filamentous cyanobacteria.

**Hormogonium formation/gas vesicle genes**

In a study of the expression of the *gvp* genes (involved in the formation of the gas vesicles) during hormogonium differentiation in *Calothrix PCC 7601*, RNA blots revealed that five transcripts hybridized with a *gvpA* probe (Damerval et al., 1987). The mapping of the 5′ and 3′ extremities, together with the use of different internal fragments of *gvpA, gvpC*, and oligonucleotides as probes showed: (i) three transcripts that all have the same 5′-end correspond to mRNAs covering either *gvpA1, gvpA1A2* or *gvpA1A2C*; (ii) a fourth transcript is a natural antisense RNA (*Fig. 1*); and (iii) a fifth mRNA species arises from *gvpD*, a non adjacent gene that is 83% homologous to *gvpA* (Csiszár et al., 1987). In addition, from time-course experiments and the degradation pattern seen on the RNA blot, a role has been proposed for the antisense RNA: since the antisense RNA can anneal with the *gvpA1,*...
gpA1A2 or gpA1A2C mRNA species, duplexes could be formed that would be substrates for putative endoribonucleases; gp mRNA species being degraded, the Gvp proteins can no longer be synthesized and the formation of the gas vesicles is efficiently stopped. Screening of the RNA blots with various gene probes has also shown that, during hormogonium differentiation, transcription of the gp genes occurs while that of the genes encoding the other genes during heterocyst differentiation in transcription of the PhepA vegetative cells that was restricted to the heterocysts, whereas light emission occurs in both cell types. Similarly, using P,,,,,,,,lzlxAB or PCC 7120. When cells were grown in air without a source of nitrogen sources used for growth have been reported for diazotrophic heterocystous strains such as Anabaena/Nostoc PCC 7120 and PCC 7937, as well as for the non-heterocystous Plectonema boryanum (Haselkorn et al., 1983; Mulligan & Haselkorn, 1989; Weiland et al., 1989; Fujita et al., 1991). To study differential gene expression during this developmental process, C. P. Wolk and co-workers have developed tools complementary to the RNA blot technique. They applied the gene fusion technology and placed the luxAB genes of Vibrio fischeri under the control of various promoter sequences. They could then precisely trace the expression of luciferase within a single cell, and have successfully used this technology with Anabaena/Nostoc species (Elhai & Wolk, 1990). On the other hand, the cat gene (encoding chloramphenicol acetyltransferase) has also been used as a reporter gene to develop a promoter-probe vector for Anabaena/Nostoc PCC 7120 (Lang & Haselkorn, 1991).

Using luxAB fusions present in trans, Elhai & Wolk (1990) have nicely demonstrated opposite patterns of transcription in the different cell types of Anabaena/Nostoc PCC 7120. When cells were grown in air without a source of combined nitrogen, light emission from P,,,,,,nifHDKluxAB was restricted to the heterocysts, whereas light emission from P,,rocLuxAB only occurred in vegetative cells. In contrast, all cells emitted light from P,,,,,,glnALuxAB fusions. These results agree with biochemical data suggesting that: (i) nitrogenase (nif) is present only in heterocyst; (ii) ribulose bisphosphate carboxylase (rbcLS) is restricted to vegetative cells; and (iii) glutamine synthetase (glnA) occurs in both cell types. Similarly, using P,,,,,,hepA luxAB or P,,,,,,hepAT7RN Apol coupled with P,,,,,,T7luxAB fusions, the transcription of the hepA (previously hetA) gene, which encodes a component of the envelope polysaccharide of heterocysts, has been shown to be confined principally to cells that display a pattern similar to that of heterocyst spacing and nif gene expression (Wolk et al., 1993).

Differential transcription has also been demonstrated for other genes during heterocyst differentiation in Anabaena/Nostoc PCC 7120. The betR gene (essential for heterocyst differentiation) is abundantly transcribed very early in the differentiation process, but only at a basal level under nitrogen-replete conditions (Buikema & Haselkorn, 1991; Black et al., 1993). A similar pattern of transcription has been observed for patS, which encodes a putative response regulator required for the differentiation of intercalary heterocysts (Li et al., 1992) and for psbA, whose expression is required for the synthesis of heterocyst glyceolipid (Buikema & Haselkorn, 1993). In contrast, the fdxH (heterocyst ferredoxin) and patAB (regulatory protein) genes are only expressed under conditions of nitrogen stepdown (Böhm & Haselkorn, 1988; Li et al., 1993).

In Anabaena/Nostoc PCC 7120, mapping of the 5’-end of mRNAs that encode the same gene product has often demonstrated multiple tsp, as for glnA, psbB, sigA and sigC (Schneider et al., 1991; Brahamsha & Haselkorn, 1991, 1992). It is not clear whether the different 5’-ends arise from multiple promoters or from specific processing since no in vitro capping experiments using the GTP transferase have been performed. Because correct transcription initiation sites have been reproduced in in vitro transcription assays for at least some of these genes, the multiple promoter hypothesis is presently favoured (Schneider et al., 1991). The underlying rationale for these multiple promoters is quite unclear, although in the case of glnA, for example, there may exist a ‘constitutive’ promoter that is always expressed and secondary promoter(s) regulated by the nitrogen status of the cells, as are the Σ42-dependent promoters of enteric bacteria (Kustu et al., 1989). For both glnA and sigC, one promoter indeed seems to be effective in transcription initiation primarily under conditions of nitrogen limitation. On the other hand, many more transcripts of the sigC gene accumulate under conditions of sulphur limitation (Brahamsha & Haselkorn, 1992). Attempts to isolate Σ42-type factors in cyanobacteria have so far been unsuccessful.

Finally, in Anabaena/Nostoc PCC 7120, as reported for Synechococcus PCC 7942 (Vega-Palas et al., 1992), nitrogen control seems to involve a global regulator, NtcA (= BifA), which belongs to the family of prokaryotic regulatory proteins that includes CRP (cAMP receptor protein), FNR, FixK, etc. BifA recognizes a sequence designated VFI binding site, ACGTTCTGAAAACGACTACAAA, which is present in the upstream regions of the rbcL, glnA and nifH genes, but also of xisA (developmentally regulated site-specific recombinase required for the rearrangement of nif genes). As proposed by Vega-Palas et al. (1992) for Synechococcus PCC 7942, NtcA could act in Anabaena/Nostoc PCC 7120 as a transcriptional activator for genes subject to ammonium repression, but it would also inhibit expression of xisA, thus acting as both a positive and a negative effector (Wei et al., 1993).

Complementary chromatic adaptation

In Calothrix PCC 7601 (Conley et al., 1988; Tandeau de Marsac et al., 1988; Federspiel & Scott, 1992) and Pseudanabaena PCC 7409 (Dubbs & Bryant, 1991) the
cpeBA and cpeCDE operons that encode the phycoerythrín subunits and their associated linker polypeptides are transcribed under green light, while transcription of the cpe2 operon that encodes the phycocyanin-2 subunits and the corresponding linkers is turned off, the converse pattern of transcription being observed under red light. In *Calothrix* PCC 7601, the adaptation mechanism not only involves controls of transcription initiation, but probably also of mRNA stability, since *de novo* protein synthesis seems to be required for an active and specific degradation of the cpe2 mRNA to occur upon a shift from red to green light (Oelmüller et al., 1989). The present working hypothesis stipulates that the light signal is sensed by a photoreversible pigment, functionally analogous to the eukaryotic phytochrome, then transduced by an unknown mechanism, and finally reaches the target genes, with some turned on and others turned off.

Very recently, protein factors which are able to bind to the cpeBA promoter region and which could behave as transcriptional effectors have been described in *Calothrix* PCC 7601 (Sobczyk et al., 1993; Schmidt-Goff & Federspiel, 1993). Sobczyk et al. (1993) reported that two proteins, RcaA and RcaB, present in partially purified fractions have DNA-binding properties towards the P_{cpeBA} sequence and that they could only be isolated from cells grown under green light. Furthermore, the affinity of RcaA, but not that of RcaB, for the DNA could be suppressed by a treatment with alkaline phosphatase implying that RcaA is active only in a phosphorylated form. The RcaA target on the DNA contains a direct repeat of an hexanucleotide (TTGTTA-N,-TTGTTA) located between nt -50 and -65. Because of its phosphoprotein nature and of the location of its binding site on P_{cpeBA}, RcaA could function like the NtrC protein which is a transcriptional activator required in its phosphoprotein nature and of the location of its binding site on P_{cpeBA}. This assumption has been proven for enteric bacteria. On the other hand, RcaB might be a specific *c* factor. Schmidt-Goff and Federspiel (1993) also described a DNA-binding protein (PebP) functionally analogous to RcaA, and which has the same target, but which is present in the cells regardless of the light wavelength employed for growth. The discrepancy between these results remains to be resolved, but may reflect differences in growth conditions (pH, temperature or growth phase).

For the cpe2 operon, a protein present only in cells grown under red light and provisionally designated RcaD exhibits affinity for the P_{cpe2} sequence. This protein loses its DNA-binding property following a treatment with alkaline phosphatase (A. Sobczyk & J. Houmard, unpublished data). On the other hand, Chiang *et al.* (1992) isolated a gene, *rcaC*, by complementation of a mutant that constitutively expressed the cpe operons and had a low level of inducible phycocyanin (cpe2 operon) even under red light. The encoded RcaC polypeptide belongs to the family of the response regulators whose activity is modulated by phosphorylation via histidine kinases, the sensors of the environmental changes. However, the sensor kinase expected to correspond to the *Calothrix* PCC 7601 RcaC remains to be found (Chiang *et al.*, 1992), as does that corresponding to the *Anabaena/Nostoc* PCC 7120 PatA regulator protein.

### Iron and copper regulation

Transcriptional controls via metal ion availability have been demonstrated in *Anabaena/Nostoc* PCC 7937. In this strain, the amount of the ferredoxin protein (PetD) and of the corresponding mRNAs present in the cells depend on iron concentrations. The steady-state level of petD mRNAs is six to ten times greater at high iron concentration (Bovy *et al.*, 1993). Part of the regulation of this petD gene has been shown to occur at the level of transcription initiation, the other parameter being the mRNA stability which is much higher when iron is present than under iron limitation. Upon shift of *Anabaena/Nostoc* PCC 7937 from low to high Cu^{2+}, there is a rapid increase in the concentration of plastocyanin (petE) mRNAs and a rapid decrease in the concentration of cytochrome f_{538} (cytA = petI) mRNAs, while no difference was observed in mRNA half-life for either gene whatever the Cu^{2+} concentration, implying that regulation occurs primarily at the level of transcription initiation. However, the changes in mRNA concentrations do not occur if chloramphenicol is added (Bovy *et al.*, 1992). This means that *de novo* protein synthesis of at least one trans-acting factor is required before the transcription of these genes is altered.

Few homologies have yet been detected between promoter sequences in a given strain, but sequence conservation is high for a given gene between different species (whether unicellular or filamentous), and the RNA polymerase of a unicellular strain may express genes from a filamentous strain. For example, insertion of heterologous sequences into the chromosome of *Synechococcus* PCC 7942 gave detectable expression of the petD, cytA (=petI) and petE genes of *Anabaena/Nostoc* PCC 7937, and of the apcE gene of *Calothrix* PCC 7601 (Bovy *et al.*, 1993; Capuano *et al.*, 1993). However, differences may exist between strains in the regulatory sequences and/or mechanisms. None of the metal-regulated *Anabaena/Nostoc* PCC 7937 genes described above retain their regulation by metal ions in *Synechococcus* PCC 7942, and even in *Anabaena/Nostoc* PCC 7120 the petE gene is not regulated by copper, while cytA expression is only slightly modulated (Bovy *et al.*, 1992, 1993).

### Conclusions and further prospects

It is increasingly evident that in all the regulatory processes under study, both transcriptional and post-transcriptional controls operate in a coordinated manner. The latter controls are probably dominant for the short-term responses to environmental changes since they may involve modifications of activities of enzymes already present in the cells or even purely chemical isomerization reactions. The reversibility of enzymic reactions largely prevents very tight controls but could permit quite subtle, rapid and sensitive responses. In contrast, transcriptional regulations allow more stringent control, but may require more time to be established. However, they are often
easier to study than the post-transcriptional regulations, since steady-state equilibria may develop when cells become adapted to the new environmental conditions and all-or-nothing responses can be observed, while it is more difficult to freeze transient modifications. Although regulatory adjustments in the expression of some genes may occur at the translational level, such controls are often superimposed on changes of the steady-state mRNA levels over longer periods. Pre-transcriptional events may sometimes also be required for gene expression, as demonstrated by the rearrangements that involve the *nif* genes during heterocyst differentiation in *Anabaena* strains (Buikema & Haselkorn, 1993).

Because of their wide distribution and their rather limited nutritional exigency, cyanobacteria are often considered as potentially useful for biotechnological engineering. However, the design of the recombinant organism carrying exogenous gene(s) will require a thorough knowledge of the conditions necessary for the expression of the gene(s) in the selected organism. In *E. coli*, *in vitro* activities of unregulated promoters may differ by several orders of magnitude. For practical purposes, the concept of promoter activity or promoter strength has been introduced. It is defined by the rate at which productively transcribing RNA polymerase leaves its start site to enter its elongation state. Each of the steps between the first contact of DNA and RNA polymerase, and elongation, can in principle be rate-limiting and therefore determine the strength of a given promoter. Best recognition *in vitro* by the RNA polymerase does not necessarily imply higher *in vivo* activity (Knaus & Bujard, 1988). These authors concluded that a stretch of about 70 bp, spanning a region between +20 and −55 bp with regards to the transcriptional start point, best defines a promoter region in *E. coli*. For cyanobacteria, the main features that have presently emerged are that: (i) promoter sequences close to that recognized by the Ex70 *E. coli* RNA polymerase would in general be expressed to a varying extent in most species; and (ii) when necessary, the regulatory sequences would preferably originate from the cyanobacterium in which the specific control is expected to take place. Indeed, as already mentioned, since transcription was sometimes much more efficient with cyanobacterial crude extracts than with purified RNA polymerase, it is likely that specific factors were missing in the purified enzyme preparation used. Empirical approaches by trial and error would thus often be required to adjust the strategy to the desired goal, and it seems premature, at present, to think of a universal means for gene expression in cyanobacteria.

The occurrence of cascades involving many intermediates is more and more documented in prokaryotic regulatory mechanisms. Similar cascades started to be recognized for signal transduction pathways in cyanobacteria. A natural antisense RNA has already been identified. The existence of anti-σ polypeptides, similar to the one recently found in *Salmonella typhimurium* (Onishi et al., 1992), and of interactions between RNA binding proteins and long leader sequences (often found 5' of cyanobacterial transcripts) will probably prove to be involved in mRNA stability and/or translational controls. Although the dissection of the many molecular mechanisms that control gene expression in cyanobacteria is essential, the major challenge for the coming years will probably be to find the links which connect the individual pathways and form the central regulatory network. Whatever the studies, it is quite important to keep in mind that the stage of growth and culture conditions may profoundly modify gene expression, that the gene copy number has to be known and that maximum caution is required in interpreting results obtained in a heterologous background.

I gratefully acknowledge N. Tandeau de Marsac, D. Campbell, M. Herdman and J. Meeks for very helpful discussions and comments on the manuscript. I also thank my collaborators A. Sobczyk and G. Schyns and all the colleagues who communicated data before publication and sent preprints and reprints.

References


Capuano, V., Thomas, J.-C., Tandeau de Marsac, N. & Houmard, J. (1993). An *in vivo* approach to define the role of the *Lcm*, the key...


Schneider, G. J., Tumer, N. E., Richaud, C., Borbély, G. &


