Plant Molecular Biology **13**: 441–456, 1989.

Isolation, sequence and transcription of the gene encoding the photosystem II chlorophyll-binding protein, CP-47, in the cyanobacterium *Anabaena* 7120

Jean D. Lang and Robert Haselkorn
Department of Molecular Genetics and Cell Biology, The University of Chicago, 920 East 58th Street, Chicago, IL 60637, USA

Received 16 May 1989; accepted 18 May 1989

Key words: amino acid sequence, *Anabaena* 7120, CP-47, cyanobacteria, DNA sequence, transcription

Abstract

The *psbB* gene from the cyanobacterium *Anabaena* 7120 was cloned and its nucleotide sequence determined. This gene codes for the photosystem II chlorophyll-binding protein CP-47. We identified an open reading frame of 1527 bases that can code for a polypeptide with a predicted molecular weight of 56254. Expression of the *psbB* gene in *Anabaena* 7120 results in two transcripts that begin 302 or 238 bp upstream of the open reading frame and both end 288 bp downstream following a 21 bp inverted repeat. We also show that the levels of these *psbB* gene transcripts as well as those of the *glnA*, *rbcLS*, and *psbA* genes, in RNA prepared from whole filaments, remain constant during heterocyst differentiation.

Introduction

CP-47 is a chlorophyll-binding protein encoded by the *psbB* gene. This gene has been cloned from the chloroplast DNA of a number of plants and from cyanobacteria, and has been shown to be very similar in sequence from all these sources [38, 59]. Early experiments have suggested that CP-47 was the protein that bound P680, the reaction center chlorophyll that initiates the primary light-induced charge separation in photosystem II (PSII) [40, 62, 63]. PSII is a protein-pigment complex the core of which consists of 5 proteins: CP-47, CP-43, D1 (the 32 kDa Q_B protein), D2 (a 34 kDa protein) and cytochrome b_{559}. This core complex is embedded in the thylakoid membrane. Recently, a PSII reaction center complex was isolated that contained only D1, D2 and cytochrome b_{559} [41]. This complex is active in the photoreversible accumulation of reduced pheophytin (the primary acceptor of electrons from P680) and exhibits the characteristic kinetics and absorption spectrum of the PSII reaction center [41].

The assembly of the PSII complex is currently being studied in the cyanobacterium *Synechocystis* 6803 [58]. *Synechocystis* 6803 is suitable for these studies because it can be genetically transformed and can grow photoheterotropically using glucose if PSII activity is missing. Inactivation of the genes encoding CP-47, CP-43 and D2 results in the loss of PSII activity. The analysis of mutants containing modified or inactivated genes for CP-47, CP-43 and D2, using antibodies against these proteins, led Vermaas *et al.* [58] to propose the following assembly scheme. D1 and D2, with

their correct prosthetic groups bound, are first inserted into the membrane in an unstable complex. CP-47 is added next to stabilize the complex while CP-43 is inserted independently of the other proteins. This model is supported by the ability to isolate active PSII complexes that contain CP-47 but lack CP-43 [62]. It has not been possible to isolate active PSII complexes that contain CP-43 but not CP-47 [41]. Thus CP-47 seems to be more closely associated with the complex than CP-43.

We describe here the isolation, sequence and transcription of the *psbB* gene in *Anabaena* 7120. *Anabaena* 7120 is a filamentous cyanobacterium that is able to fix nitrogen as well as to carry out oxygenic photosynthesis. These two processes are not compatible since the enzyme involved in nitrogen fixation, nitrogenase, is sensitive to the oxygen evolved during photosynthesis. To overcome this problem, *Anabaena* 7120 differentiates specialized cells called heterocysts at regular intervals along the filament in response to nitrogen starvation. Heterocysts develop from vegetative

Table 1. Cellular location of *Anabaena* gene products.

Gene	Product	Product location	Reference
glnA	glutamine synthetase	V/H	57, 6
nif HDK operon			
nifH	nitrogenase reductase	/H	37, 7, 21
nifD	α subunit of dinitrogenase	/H	31, 7, 21
nifK	β subunit of dinitrogenase	/H	35, 7, 21
rbcLS	large and small subunits of ribulose-1,5-bisphosphate carboxylase	V/	16, 43, 1, 8, 14, 60
psbA	D1 (the 32 kDa Q_B protein)	V/H	17, 11
atpB/E	β, ε subunits of ATPase	V/?	15
Phycocyanin operon			
cpcB	β subunit of phycocyanin	V/?	4, 22, 54, 61, 64
cpcA	α subunit of phycocyanin	V/?	4, 22, 54, 61, 64
cpcC	rod linker	V/?	4, 22, 54, 61, 64
cpcD	rod-capping protein	V/?	4, 22, 54, 61, 64
cpcE	unknown	V/?	4, 22, 54, 61, 64
cpcF	unknown	V/?	4, 22, 54, 61, 64
woxA	33 kDa protein of the PSII associated water splitting complex	V/?	D. Borthakur
nifB operon			
nifB	required for Fe-Mo cofactor synthesis	/H	M. Mulligan
fdxN	ferredoxin like protein	/H	39
nifS	involved in maturation of nitrogenase complex	/H	M. Mulligan
nifU	involved in maturation of nitrogenase complex	/H	M. Mulligan
petF	vegetative cell ferredoxin	V/	2, 13
fdxH	heterocyst ferredoxin	/H	9, 49

cells, but differ from them in several ways: heterocysts are surrounded by a double-layered glycolipid-containing envelope, they have lost phycobiliproteins and PSII activity, their levels of enzymes of the oxidative pentose pathway are elevated, and they do not fix CO_2 [27]. These changes create an anaerobic environment that allows nitrogen fixation.

One of the goals in our laboratory is to determine the molecular mechanisms that regulate genes during heterocyst differentiation. Previous experiments have shown that transcripts for the *glnA* and *nifH* genes (see Table 1) are present during the induction of nitrogenase while those for *rbcLS* are absent [28]. These changes are correlated with parallel changes in the corresponding enzyme activity during induction of nitrogenase. The activity of glutamine synthetase increases slightly and that of ribulose-1,5-bisphosphate carboxylase disappears [45, 1]. We reexamined the levels of transcripts of *glnA*, *nifH* and *rbcLS* and studied the transcription of *psbB* and *psbA* during heterocyst differentiation.

Another goal is to define the sequences that make up a promoter used during vegetative growth. We have mapped the 5′ ends of transcripts of the *psbB* gene, tested appropriate sequences for promoter activity using RNA polymerase *in vitro* and then compared these sequences with proposed promoter sequences of other vegetative cell genes, in an effort to arrive at an *Anabaena* 7120 promoter consensus sequence.

Materials and methods

Materials

Restriction endonucleases, T4 polynucleotide kinase, T4 DNA ligase, alkaline phosphatase and DNA 3′ End Labeling Kit were obtained from Boehringer Mannheim and Bethesda Research Laboratories. AMV reverse transcriptase was obtained from BioRad. The M13 universal sequencing primer, Klenow fragment of DNA polymerase I, deoxynucleoside and dideoxynucleoside triphosphates were obtained from Pharmacia. A DNA sequencing kit containing modified bacteriophage T7 DNA polymerase [53] was purchased from United States Biochemical Corporation. [γ-^{32}P]dATP (3000 Ci/mmol), [α-^{32}P]-dCTP (3000 Ci/mmol), [^{35}S]dATPαS (>600 Ci/mmol) and CoTP[α-^{32}P] (5000 Ci/mmol) were obtained from Amersham and NEN Research Products. Oligonucleotides were synthesized on an Applied Biosystems machine by Paul Gardner.

Strains and culture conditions

Anabaena 7120 was grown in Kratz and Myers [29] medium supplemented with 2.5 mM $(NH_4)_2SO_4$ as indicated. Cultures were bubbled with 1% CO_2 in air and grown under fluorescent light at an initial light intensity of 140 $\mu E/m^2 \cdot s$ for 24 h and then at 440 $\mu E/m^2 \cdot s$ for 72 h. *Escherichia coli* MC1061 [12] was used to maintain the plasmids pBR322 and pBR328 [10]. *E. coli* JM105 [65] was used to propagate M13 clones, and *E. coli* LE392 (obtained from R. Davis) was used as host for recombinant bacteriophage lambda.

Isolation of DNA and RNA, Southern and Northern hybridizations

High molecular weight DNA was prepared from *Anabaena* 7120 as described [36]. RNA was isolated as described [26]. Bacteriophage and plasmid DNA were isolated as described [32]. Southern blots and plaque screens were done as described [5, 32]. For preparation of Northern blots, RNA was glyoxalated [55] and fractionated by agarose gel electrophoresis. RNA was transferred to nitrocellulose in 20× SSPE [32]. All blots were hybridized overnight in 6× SSPE, 0.5% sodium dodecyl sulfate (SDS), and 5× Denhardt's solution with probes that were nick-translated [32]. Hybridizations with the spinach *psbB* gene as a probe were done at 58 °C; all other hybridizations were done at 65 °C. All blots were washed in 0.5 × SSPE and 0.2% SDS

for 1 h. Filters probed with the spinach *psbB* gene were washed at 58 °C; all others were washed at 65 °C.

Isolation of psbB *gene*

A library of *Hin*d III fragments of *Anabaena* 7120 DNA cloned into the bacteriophage lambda vector NM762 [46] was screened by plaque hybridization using the cloned spinach *psbB* gene [38] as a probe. The probe was prepared by isolating the 1.5 kb *Eco* RI-*Sal* I fragment from pWHsp 207 [38]. This fragment contained 1 kb of coding sequence for spinach CP-47. Three separate screenings using, in addition, a library of *Eco* RI fragments of *Anabaena* 7120 DNA cloned into the lambda vector Charon 4 [46] were needed to isolate overlapping fragments. The plasmids pBR322 and pBR328 were used for subcloning fragments from the phage. The translation stop codon and downstream sequences were cloned by constructing a size-selected library of *Anabaena* 7120 DNA in pBR322. This library was screened by colony hybridization [32].

Sequence analysis

The DNA sequence was obtained by using the methods of Maxam and Gilbert [34] and Sanger [47]. For Sanger sequence analysis, DNA fragments were cloned into the M13 vectors MP18 and MP19 [44]. The sequence was determined with the universal primer (Pharmacia) and specific synthetic oligonucleotides as primers.

The DNA sequence obtained for the *Anabaena* 7120 *psbB* gene was analyzed using UWGCG software [20]. Possible secondary structures downstream of *psbB* were identified using the Fold program [66]. Free energy of formation for these structures was also calculated using the Fold program. The *psbB* genes from *Anabaena* 7120, spinach, and *Synechocystis* 6803 were compared using the Gap program [20]. Hydropathy profiles for the derived amino acid sequences of CP-47 proteins from *Anabaena* 7120, spinach, and *Synechocystis* 6803 were calculated with the method of Kyte and Doolittle [30]. Plots were generated with the DNA Strider program [33].

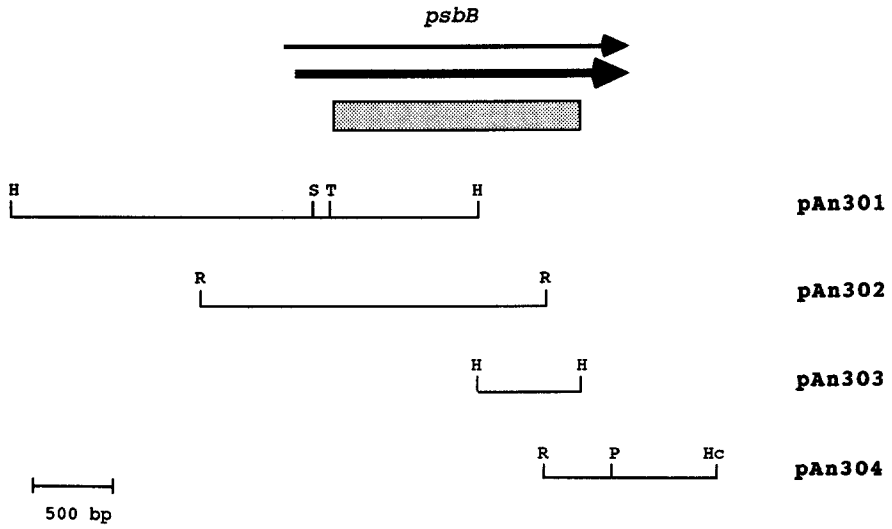


Fig. 1. Restriction map of the region containing the *Anabaena* 7120 *psbB* gene. pAn301-3 contain fragments cloned from lambda libraries. pAn304 contains a fragment isolated from a size-selected plasmid library. Arrows indicate the two transcripts for the gene that initiate at 238 bp or 302 bp upstream and terminate 288 bp downstream from the coding region. The thicker arrow indicates the most abundant transcript. Shaded box represents the open reading frame. H, *Hin*d III; Hc, *Hin*c II; P, *Pvu* I; S, *Sau*3 AI; T, *Taq* I.

S1 nuclease and primer extension analyses

For S1 protection of the 5′ ends of the *psbB* transcript, the 810 bp *Eco* RI-*Taq* I fragment contained in pAn302 (Fig. 1) was used. The plasmid was end-labeled with polynucleotide kinase at the *Taq* I site, cut with *Eco* RI and the resulting 810 bp fragment isolated. This fragment was denatured and annealed with 35 μg of *Anabaena* 7120 RNA or yeast tRNA and digested with S1 nuclease [57]. The Maxam and Gilbert G + A sequencing reaction was run using the same end-labeled fragment in order to generate a sequencing ladder.

The 5′ ends of the transcripts were confirmed by primer extension. The labeled fragment described above was cut with *Sau*3 AI (Fig. 1) and the resulting 100 bp fragment isolated. This fragment was denatured and annealed with 35 μg of *Anabaena* 7120 RNA and extended using reverse transcriptase as previously described [25]. The sequencing ladder was the same as the one used in S1 mapping.

S1 nuclease protection was also used to map the 3′ end of the transcripts. Two fragments from pAn304 (Fig. 1) were used; one containing the 820 bp *Hin*d III-*Hin*c II fragment and the other containing the 650 bp *Pvu* I-*Hin*c II fragment. The 820 bp *Hin*d III-*Hin*c II fragment was used to roughly determine the 3′ end (data not shown). The *Hin*d III and *Pvu* I sites were labeled with cordycepin triphosphate and terminal transferase as suggested by the enzyme manufacturer (Boehringer Mannheim). The DNA was then cut with *Bam* HI and the resulting labeled fragments were isolated (*Bam* HI cuts within the vector, 190 bp from the insert *Hin*c II site). Each fragment was then denatured and annealed with 35 μg of *Anabaena* 7120 or yeast tRNA and digested with S1 nuclease. The reactions were done both at room temperature and at 37 °C. The reactions at room temperature were done to avoid digestion of duplex molecules in AT-rich regions. At 37 °C, these regions can 'breathe' thus creating single-stranded regions that are susceptible to S1 nuclease. Maxam and Gilbert sequencing reactions (G, G + A, C, C + T) were run using the *Pvu* I-*Bam* HI fragment in order to generate a sequencing ladder.

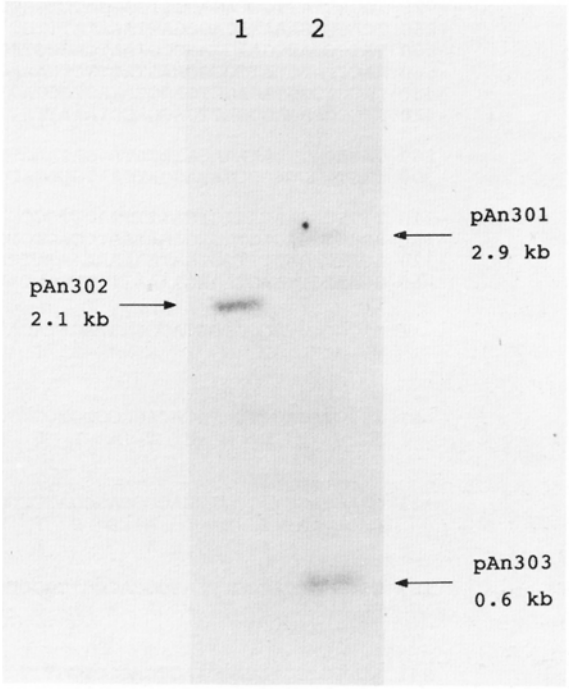
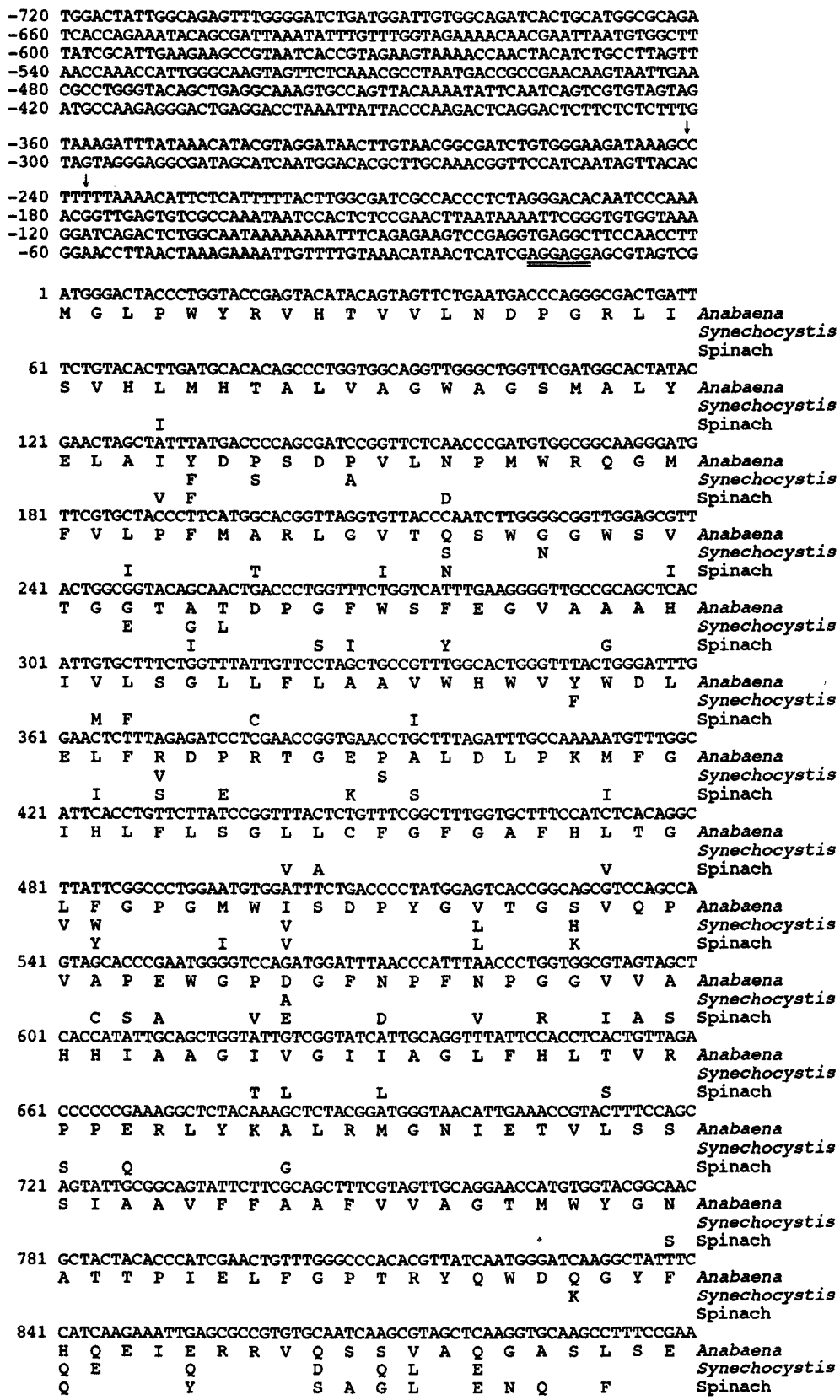


Fig. 2. Hybridization of the spinach *psbB* gene to total *Anabaena* 7120 DNA (4 μg) digested with *Eco* RI (lane 1) or *Hin*d III (lane 2). The fragment used as a probe was isolated from pWHsp 207 and contained 1 kb of coding sequence for spinach CP-47 [38]. Fragments labeled pAn301–3 were cloned from lambda libraries. *Hin*d III-digested lambda DNA was used as a size standard.

Results

Isolation of the psbB *gene*

Southern blots of total *Anabaena* 7120 DNA digested with *Eco* RI or *Hin*d III were probed with a fragment containing the spinach *psbB* gene [38] under conditions of low stringency. One *Eco* RI fragment of 2.1 kb and two *Hin*d III fragments of 2.9 kb and 0.6 kb were identified by the probe (Fig. 2). The sizes and numbers of these fragments indicated that the gene is single-copy in *Anabaena* 7120. To clone the gene, various cosmid and recombinant lambda libraries were screened. The gene was not found in any of the


-720 TGGACTATTGGCAGAGTTTGGGGATCTGATGGATTGTGGCAGATCACTGCATGGCGCAGA
-660 TCACCAGAAATACAGCGATTAAATATTTGTTTGGTAGAAAACAACGAATTAATGTGGCTT
-600 TATCGCATTGAAGAAGCCGTAATCACCGTAGAAGTAAAACCAACTACATCTGCCTTAGTT
-540 AACCAAACCATTGGGCAAGTAGTTCTCAAACGCCTAATGACCGCCGAACAAGTAATTGAA
-480 CGCCTGGGTACAGCTGAGGCAAAGTGCCAGTTACAAAATATTCAATCAGTCGTGTAGTAG
-420 ATGCCAAGAGGGACTGAGGACCTAAATTATTACCCAAGACTCAGGACTCTTCTCTCTTTG
-360 TAAAGATTTATAAACATACGTAGGATAACTTGTAACGGCGATCTGTGGGAAGATAAAGCC
-300 TAGTAGGGAGGCGATAGCATCAATGGACACGCTTGCAAACGGTTCCATCAATAGTTACAC
-240 TTTTTAAAACATTCTCATTTTTACTTGGCGATCGCCACCCTCTAGGGACACAATCCCAAA
-180 ACGGTTGAGTGTCGCCAAATAATCCACTCTCCGAACTTAATAAAATTCGGGTGTGGTAAA
-120 GGATCAGACTCTGGCAATAAAAAAAATTTCAGAGAAGTCCGAGGTGAGGCTTCCAACCTT
-60 GGAACCTTAACTAAAGAAAATTGTTTTGTAAACATAACTCATCGAGGAGGAGCGTAGTCG
1 ATGGGACTACCCTGGTACCGAGTACATACAGTAGTTCTGAATGACCCAGGGCGACTGATT
M G L P W Y R V H T V V L N D P G R L I Anabaena
Synechocystis
Spinach
61 TCTGTACACTTGATGCACACAGCCCTGGTGGCAGGTTGGGCTGGTTCGATGGCACTATAC
S V H L M H T A L V A G W A G S M A L Y Anabaena
Synechocystis
I Spinach
121 GAACTAGCTATTTATGACCCCAGCGATCCGGTTCTCAACCCGATGTGGCGGCAAGGGATG
E L A I Y D P S D P V L N P M W R Q G M Anabaena
F S A Synechocystis
V F D Spinach
181 TTCGTGCTACCCTTCATGGCACGGTTAGGTGTTACCCAATCTTGGGGCGGTTGGAGCGTT
F V L P F M A R L G V T Q S W G G W S V Anabaena
S N Synechocystis
I T I N I Spinach
241 ACTGGCGGTACAGCAACTGACCCTGGTTTCTGGTCATTTGAAGGGGTTGCCGCAGCTCAC
T G G T A T D P G F W S F E G V A A A H Anabaena
E G L Synechocystis
I S I Y G Spinach
301 ATTGTGCTTTCTGGTTTATTGTTCCTAGCTGCCGTTTGGCACTGGGTTTACTGGGATTTG
I V L S G L L F L A A V W H W V Y W D L Anabaena
F Synechocystis
M F C I Spinach
361 GAACTCTTTAGAGATCCTCGAACCGGTGAACCTGCTTTAGATTTGCCAAAAATGTTTGGC
E L F R D P R T G E P A L D L P K M F G Anabaena
V S Synechocystis
I S E K S I Spinach
421 ATTCACCTGTTCTTATCCGGTTTACTCTGTTTCGGCTTTGGTGCTTTCCATCTCACAGGC
I H L F L S G L L C F G F G A F H L T G Anabaena
Synechocystis
V A V Spinach
481 TTATTCGGCCCTGGAATGTGGATTTCTGACCCCTATGGAGTCACCGGCAGCGTCCAGCCA
L F G P G M W I S D P Y G V T G S V Q P Anabaena
V W V L H Synechocystis
Y I V L K Spinach
541 GTAGCACCCGAATGGGGTCCAGATGGATTTAACCCATTTAACCCTGGTGGCGTAGTAGCT
V A P E W G P D G F N P F N P G G V V A Anabaena
A Synechocystis
C S A V E D V R I A S Spinach
601 CACCATATTGCAGCTGGTATTGTCGGTATCATTGCAGGTTTATTCCACCTCACTGTTAGA
H H I A A G I V G I I A G L F H L T V R Anabaena
Synechocystis
T L L S Spinach
661 CCCCCCGAAAGGCTCTACAAAGCTCTACGGATGGGTAACATTGAAACCGTACTTTCCAGC
P P E R L Y K A L R M G N I E T V L S S Anabaena
Synechocystis
S Q G Spinach
721 AGTATTGCGGCAGTATTCTTCGCAGCTTTCGTAGTTGCAGGAACCATGTGGTACGGCAAC
S I A A V F F A A F V V A G T M W Y G N Anabaena
Synechocystis
S Spinach
781 GCTACTACACCCATCGAACTGTTTGGGCCCACACGTTATCAATGGGATCAAGGCTATTTC
A T T P I E L F G P T R Y Q W D Q G Y F Anabaena
K Synechocystis
Spinach
841 CATCAAGAAATTGAGCGCCGTGTGCAATCAAGCGTAGCTCAAGGTGCAAGCCTTTCCGAA
H Q E I E R R V Q S S V A Q G A S L S E Anabaena
Q E Q D Q L E Synechocystis
Q Y S A G L E N Q F Spinach

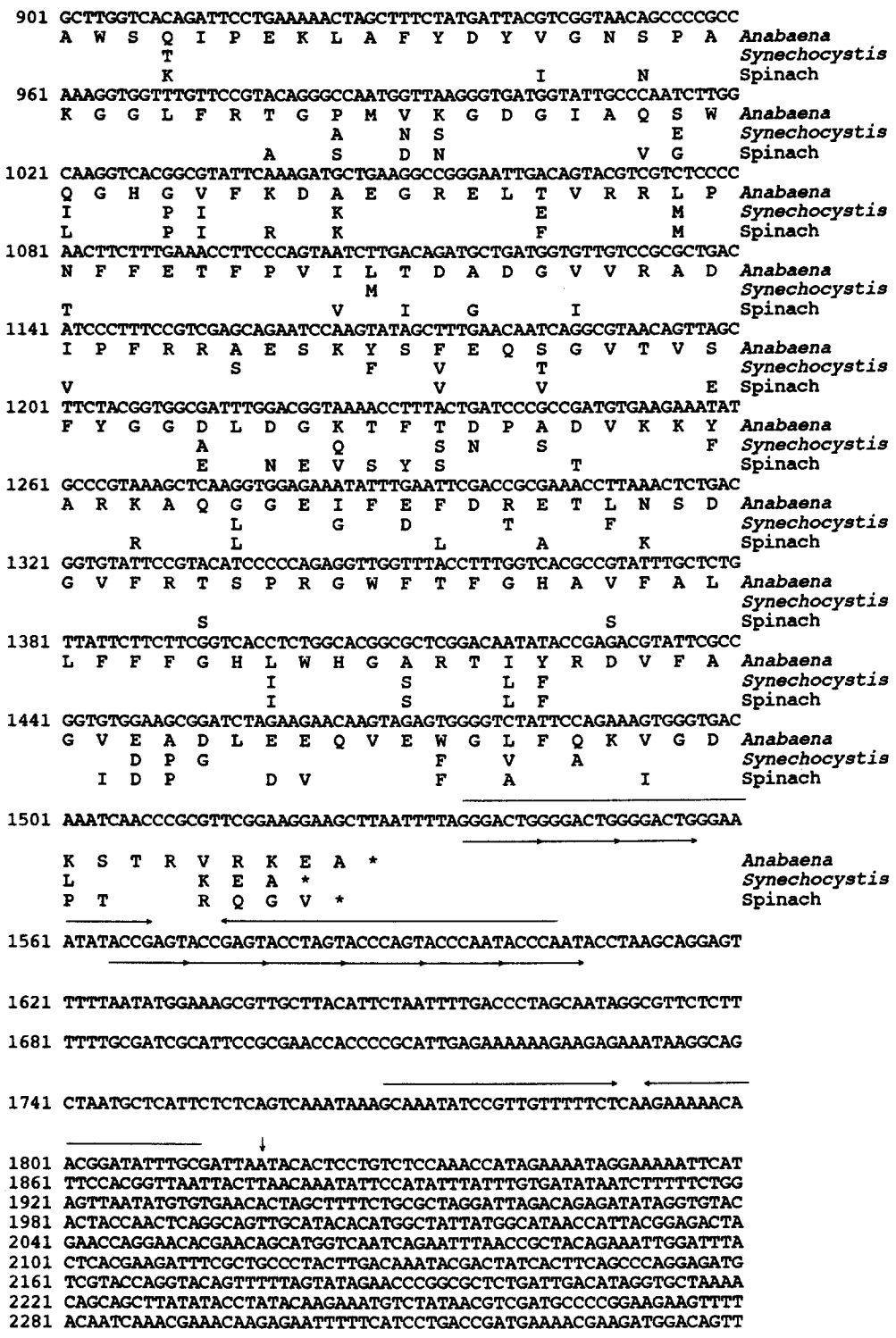


Fig. 3. Nucleotide sequence of the noncoding strand of the *psbB* gene from *Anabaena* 7120. The derived amino acid sequence of CP-47 of *Anabaena* 7120 is compared to the CP-47 sequences of *Synechocystis* 6803 and spinach [59, 38]. Only differences are shown. Vertical arrows indicate the two transcription initiation sites and putative transcription termination site. An *E. coli*-like Shine-Dalgarno sequence is underlined [50]. Direct repeats are underlined with short arrows. Sequences that may form stable stem and loop secondary structures are overlined with long arrows. Numbering of nucleotides begins with the first A of the first methionine codon.

cosmid libraries, but most of the gene was isolated on three overlapping fragments cloned from the lambda libraries (Fig. 1).

By comparing the sequence obtained from these fragments to the published sequence for the *psbB* gene from spinach [38], it was determined that the stop codon was missing from the 0.6 kb *Hin*d III fragment thought to contain the 3′ end of the gene. To clone the stop codon and downstream sequences, a size-selected library was constructed. Fragments of approximately 1.0 kb were isolated from an *Eco* RI-*Hin*c II digest of *Anabaena* 7120 DNA and cloned into the *Eco* RI/*Eco* RV sites of pBR322. These constructs were used to transform *E. coli* MC1061 and the resulting colonies were screened with the 240 bp *Eco* RI-*Hin*d III fragment isolated from pAn303. A plasmid containing a 1.05 kb *Eco* RI-*Hin*c II fragment that overlapped with the 0.6 kb *Hin*d III fragment was thus identified.

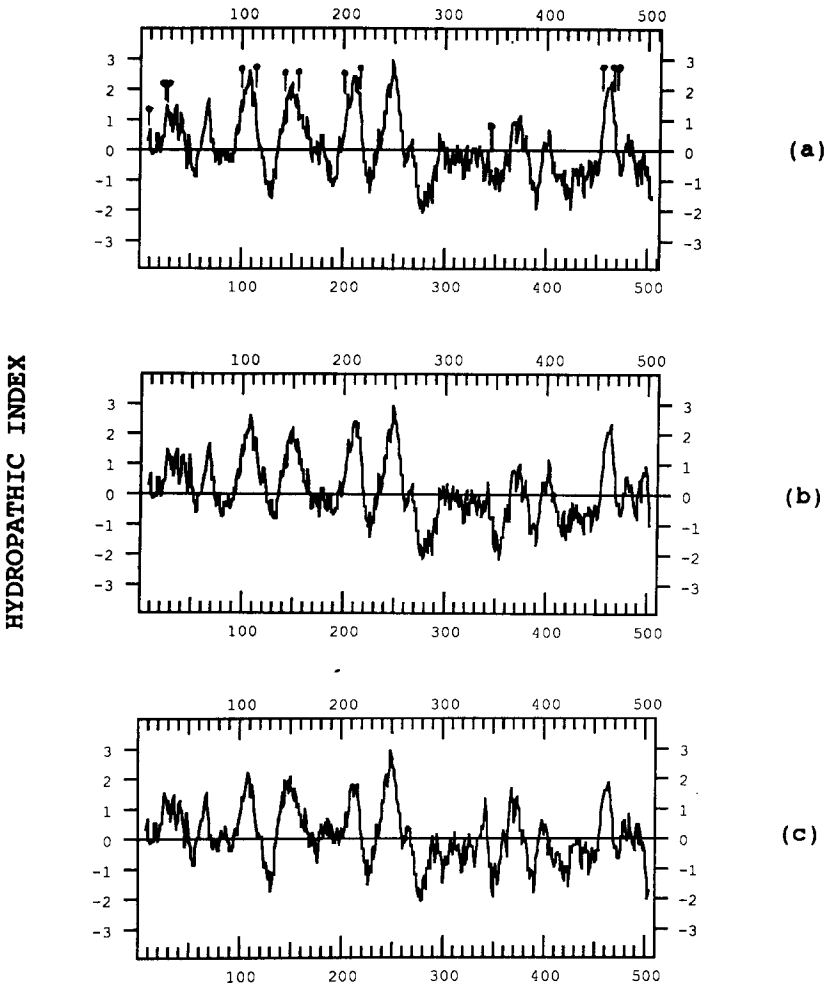


Fig. 4. Hydropathy profiles of the amino acid sequence derived from the *psbB* genes from *Anabaena* 7120 (a), *Synechocystis* 6803 (b) and spinach (c). Amino acid number is plotted against the average hydrophobicity of eleven surrounding residues according to Kyte and Doolittle [30]. Vertical symbols indicate the fourteen conserved histidine residues at positions 9, 23, 26, 100, 114, 142, 157, 201, 202, 216, 343, 455, 466 and 469.

DNA sequence analysis

The sequence obtained from the 2.1 kb *Eco* RI fragment showed that a base was inserted at position 44 (amino acid 15) causing a shift in the reading frame and termination at amino acid 46. Other cyanobacterial PSII genes are known to be lethal in *E. coli* [25, 59]. Since the 2.1 kb fragment contained most of the *psbB* open reading frame (ORF), it seemed possible that this gene product is also lethal to *E. coli* and that a mutation occurred and was selected during subcloning from lambda into the plasmid. To test this possibility, a fragment containing only the first 130 bases of the ORF was recloned from lambda and sequenced. This fragment did not contain the inserted base. Toxicity of the *psbB* gene product probably accounts for our inability to find the gene in cosmid libraries.

The final nucleotide sequence obtained from the various cloned fragments is shown in Fig. 3. Examination of the sequence upstream of the *Anabaena* 7120 *psbB* gene coding region shows an *E. coli*-like ribosome binding site [50], AGGAGG, beginning 11 bp upstream of the putative start codon. An ORF of 1527 bases is identified that can code for a polypeptide with a predicted M_r of 56254. The ORF for the *Anabaena* 7120 *psbB* gene is very similar to the ORFs of 1524 and 1521 bases determined for the spinach and *Synechocystis* 6803 *psbB* genes [38, 59]. The *Anabaena* 7120 *psbB* gene DNA sequence is 69% similar to that of the spinach gene DNA sequence and 76% similar to that of the *Synechocystis* 6803 gene DNA sequence.

A comparison of the translated ORFs revealed that the spinach protein sequence is 88% similar and the *Synechocystis* 6803 protein sequence is 93% similar to the *Anabaena* 7120 protein sequence. The hydropathy plots for the three proteins (Fig. 4) are almost identical indicating that they may be folded the same way in the thylakoid membrane. We defer until the Discussion the possible significance of the conserved histidines shown in Fig. 4a.

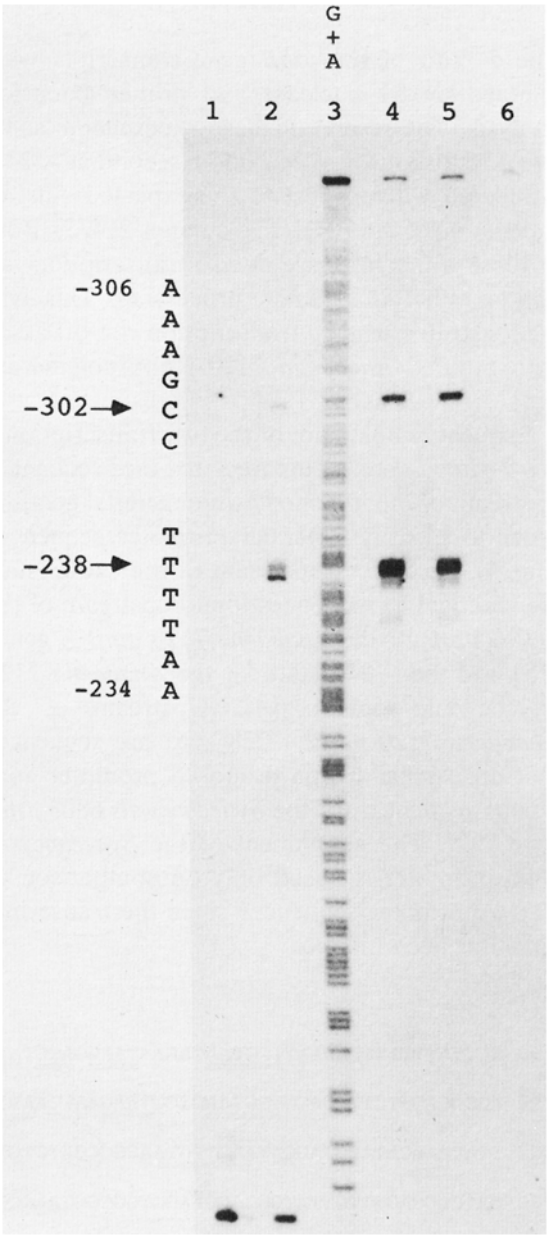


Fig. 5. Identification of the start sites for *Anabaena* 7120 *psbB* transcription. A 5′ end-labeled *Eco* RI-*Taq* I fragment (Fig. 1) containing sequences upstream of the *psbB* gene was hybridized to *Anabaena* 7120 RNA (lane 4, 5) or yeast tRNA (lane 6) and then digested with S1 nuclease. The *Eco* RI-*Taq* I fragment was also used to generate a sequencing ladder (lane 3). The S1 result was confirmed by extension of a primer derived from the *Eco* RI-*Taq* I fragment (lane 2). Lane 1 contains the unextended primer. Arrows mark putative transcription start sites. Sequence numbering follows Fig. 3.

Transcription of psbB

The 5′ end of the *psbB* gene transcripts were mapped by S1 nuclease and primer extension analysis. The results of these experiments show two 5′ ends, one at −302 and a second at −238, numbered with respect to the expected start of translation of the ORF (Fig. 5, lanes 2, 4, 5). Both of these transcripts are due to transcription initiation rather than RNA processing. This was demonstrated using a transcription run-off assay with purified *Anabaena* 7120 RNA polymerase [48].

Sequences upstream of the two transcriptional start sites were compared to the sequences upstream of other cloned cyanobacterial genes in order to identify potential promoter sequences (Fig. 6). Sequences upstream of the −302 start are identical to sequences found upstream of the −196 start for the *Anabaena* 7120 *atpB/E* genes [15] and the −243 start for the *Anabaena* 7120 phycocyanin operon [4]. Upstream of the *Anabaena* 7120 *psbB* −238 start are sequences that are similar to the proposed promoter consensus sequences for the *Synechocystis* 6803 *psbB* gene [59]. The assignment of the *Synechocystis* 6803 promoter is based only on comparison to *E. coli* promoter sequences since the transcripts have not been mapped.

The 3′ end of the *psbB* transcripts were also mapped using S1 nuclease analysis and found to end 288 bp downstream of the expected stop codon (Fig. 7, lanes 3, 4). Beginning 49 bp before the end of the transcript there is a 21 bp inverted repeat capable of forming a stem and loop structure with a predicted free energy of formation of −100 kJ/mol (−24 kcal/mol) (Fig. 8b), which appears to function as a termination or processing site. In addition to the cleavage 6 bp after this 21 bp inverted repeat, two weaker cleavages were seen in the 3′ end determinations with S1 nuclease. One cleavage occurred in the expected loop of the structure that could be formed by the 21 bp inverted repeats. This cleavage may be due to the formation of a cruciform structure in the DNA since cleavage occurred when the S1 nuclease reactions were carried out at 37 °C (Fig. 7, lane 3), at room temperature (Fig. 7, lane 4) and also when the S1 nuclease reactions were carried out using yeast tRNA instead of *Anabaena* 7120 RNA (Fig. 7, lane 2). The other additional cleavages occurred in the first half of the 21 bp inverted repeat. These cleavages may be due to digestion of the DNA-RNA duplex molecules in an AT-rich region that breathes at 37 °C (Fig. 7, lane 3), but not at room temperature (Fig. 7, lane 4).

Starting 6 bp after the translation stop codon

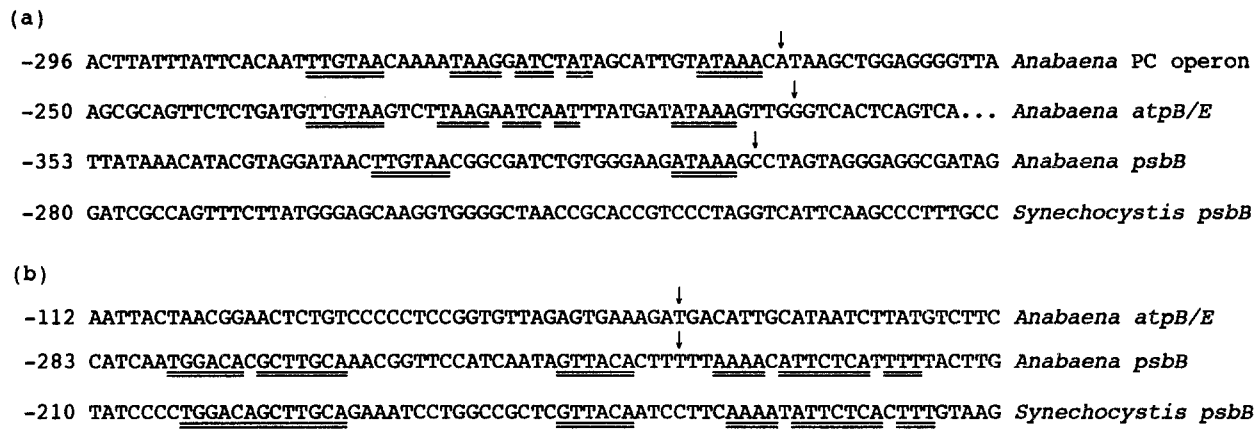


Fig. 6. Comparison of sequences upstream of the *Anabaena* 7120 *psbB* −302 start (a) or −238 start (b) with sequences upstream of the *Synechocystis* 6803 *psbB* gene [59], *Anabaena* 7120 *atpB/E* genes [15] and *Anabaena* 7120 phycocyanin (PC) operon [4]. Arrows mark putative transcription start sites determined by S1 nuclease protection and/or primer extension. Sequences shared by more than one gene are underlined. Numbering begins with −1 representing the first nucleotide before the open reading frame.

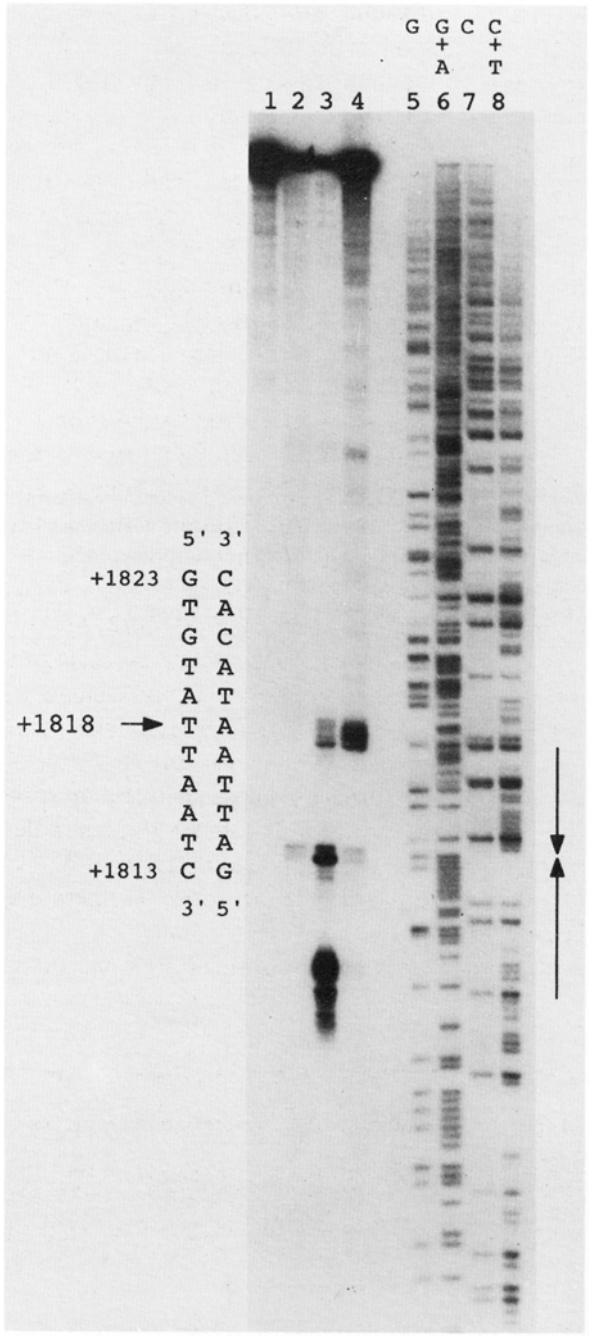


Fig. 7. Identification of the 3′ end of the transcripts of the *Anabaena* 7120 *psbB* gene. A 3′ end-labeled *Pvu* I-*Hinc* II fragment (Fig. 1) containing sequences downstream of the *psbB* gene was hybridized to *Anabaena* 7120 RNA (lanes 3, 4) or yeast tRNA (lane 2) and digested with S1 nuclease. The *Pvu* I-*Hinc* II fragment was also used to generate sequencing tracks (lanes 5, 6, 7, 8). S1 reactions were done at two temperatures: room temperature (lanes 2, 4) and at 37 °C (lane 3). Lane 1 contains the starting fragment. Horizontal arrow marks the putative transcription termination site. Sequence numbering follows Fig. 3. Vertical arrows mark the 21 bp inverted repeat.

there is a set of 7 bp repeats that could also form a stem and loop structure with a predicted free energy of formation of − 112 kJ/mol (− 26.8 kcal/mol) (Fig. 8a). However, the results of the 3′ end determinations with S1 nuclease show that this potential stem and loop structure is neither a transcription termination site nor an RNA processing site since there are no transcripts ending right after these repeats.

Transcripts of the *psbB* gene were next studied using Northern blots. A probe containing the *Anabaena* 7120 *psbB* gene was hybridized to total *Anabaena* 7120 RNA isolated from cells grown on ammonia (N +) or induced to differentiate heterocysts by suspension in medium lacking combined nitrogen (N −). Fig. 9a shows the result of probing a blot containing RNA that was isolated from a differentiating culture with a fragment of DNA containing *psbB*. Samples were removed from a differentiating culture at six hour intervals and RNA was isolated from them (RNA was isolated by Jim Golden [24]). One cell generation is necessary to complete differentiation and transcripts for nitrogenase can be found late in this process at about 24 h. One transcript of 2.1 kb hybridized to the *psbB* probe, in agreement with the size predicted by S1 nuclease mapping of the 5′ and 3′ ends. The level of message for the *psbB* gene is approximately constant during the induction period. Similar patterns of transcription were seen for *rbcLS*, *glnA* and *psbA* (Fig. 9b, c , d).

Discussion

The *Anabaena* 7120 *psbB* gene has been cloned based on similarity to the spinach *psbB* gene. Translation of the observed nucleotide sequence shows an ORF of 509 amino acids that is very similar to the ORFs of 508 amino acids and 507 amino acids for CP-47 from spinach and *Synechocystis* 6803. Analysis of the *Anabaena*

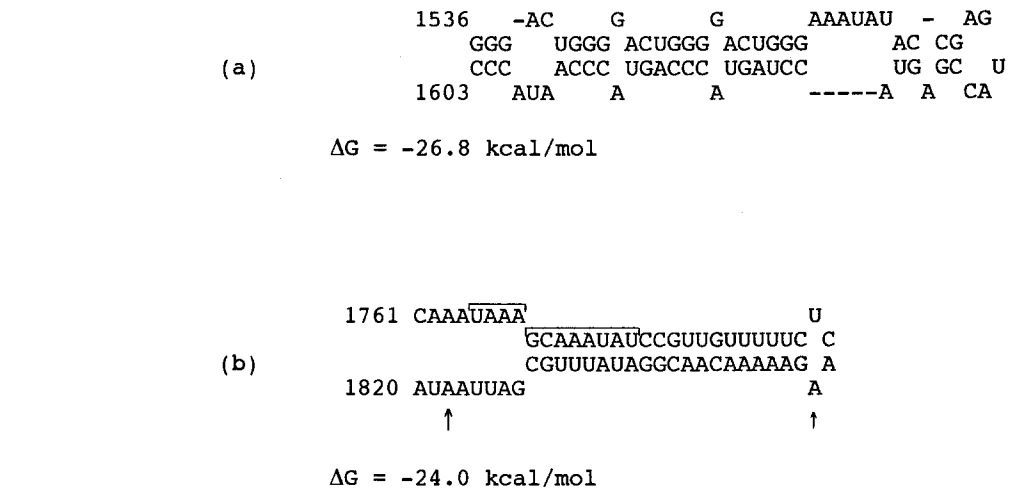


Fig. 8. Secondary structures that may form downstream of the *Anabaena* 7120 *psbB* gene. a. The proposed secondary structure beginning 6 bp downstream from the *psbB* translational termination codon. b. The proposed secondary structure beginning 49 bp before the end of the *psbB* transcripts. Numbering follows Fig. 3. Large arrow indicates the putative transcription termination site. Brackets indicate additional S1 nuclease cleavages that occurred only when the reactions were carried out at 37 °C (Fig. 7, lane 3). Small arrow indicates additional S1 nuclease cleavages that occurred when reactions were carried out at 37 °C (Fig. 7, lane 3), room temperature (Fig. 7, lane 4) and also when the S1 nuclease reactions were carried out using yeast tRNA instead of *Anabaena* 7120 RNA (Fig. 7, lane 2).

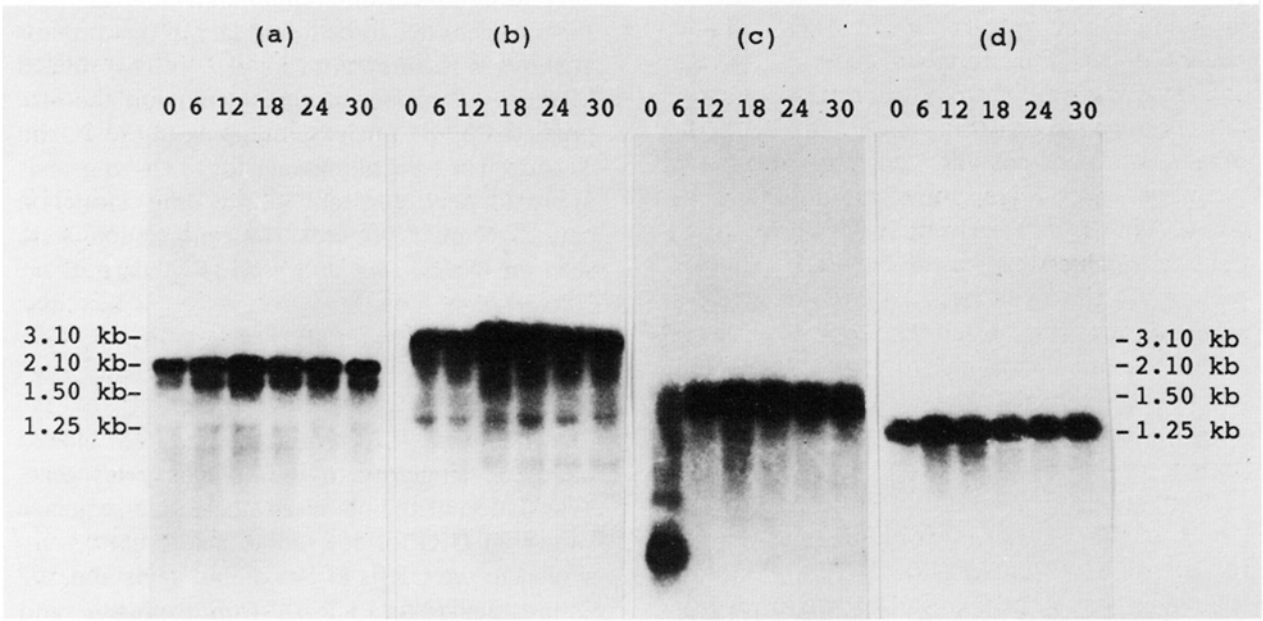


Fig. 9. Northern blots to detect transcripts in *Anabaena* 7120. Cells grown on ammonia (N+) were induced to differentiate heterocysts by suspension in medium lacking combined nitrogen (N−). RNA was isolated from cells at six-hour intervals (0, 6, 12, 18, 24, 30 h after induction of heterocyst differentiation). The blots were hybridized with DNA fragments containing the following *Anabaena* 7120 genes. a. *psbB* (isolated from pAn301). b. *rbcS* (isolated from pAn606 [43]). c. *glnA* (isolated from pAn503 [57]). d. *psbA* (isolated from pAn625 [17]). Numbers to the left and right indicate the transcript sizes as determined using RNA markers (Bethesda Research Laboratories). Mature heterocysts are first seen at 18 h, at which time *nif* gene mRNA is first detected [24]. The apparent selective degradation of *glnA* mRNA in the 0 time sample is not understood.

7120 protein sequence shows features that have been described previously for the spinach and *Synechocystis* 6803 protein sequences. For example, the hydropathy plots for all three proteins predict six membrane spanning regions. Fourteen of the histidines found in the spinach and *Synechocystis* 6803 proteins are also found in the *Anabaena* 7120 protein (Fig. 4). Of these, four pairs are located precisely in the center of putative membrane-spanning regions. The spacing between paired histidines is 14, 15, 14 or 15 and 11 or 14 residues, respectively. These histidine residues may be involved in chlorophyll binding. An X-ray crystallographic analysis of the reaction center from the purple bacterium *Rhodopseudomonas viridis* has determined that the four bacteriochlorophyll *b* associated with the reaction center are liganded to histidines [18, 19]. However, since 40 chlorophyll molecules are associated with CP-47 and CP-43, a second PSII chlorophyll-binding protein, and these two proteins contain a total of only 27 conserved histidine residues, additional amino acids may be involved in binding chlorophyll [59, 63]. The putative membrane-spanning region centered at residue 250 lacks histidine entirely. The residues at positions corresponding to the paired histidines in the other membrane-spanning regions are, here, methionine and serine. Something other than an amino acid may bind chlorophyll as was found in the bacteriochlorophyll *a* protein from the green photosynthetic bacterium *Prosthecochloris aestuarii* [56]. This bacteriochlorophyll *a* protein binds 7 bacteriochlorophyll *a*, 5 to histidines, one to a bound water molecule and one to a carbonyl oxygen from the polypeptide backbone of the protein.

Analyses of sequences upstream and downstream of the *Anabaena* 7120 *psbB* gene address the question of transcriptional control of this gene. The first surprising finding was of two 5′ ends at −302 and −238. *In vitro* transcription using purified RNA polymerase shows that both of these ends are the result of transcription initiation and that the promoter before the −238 start is the stronger of the two [48]. Multiple promoters have also been found for the *glnA* gene (encoding glutamine synthetase), which is transcribed from different promoters during growth utilizing ammonia or fixed nitrogen [57]. Two 5′ ends have recently been found for the *Anabaena* 7120 *atpB/E* genes with the shorter transcripts being the most abundant [15] as is the case for *psbB*. The reason the *psbB* gene has two promoters is unknown, but it is not a general characteristic of genes encoding proteins that are associated with PSII since *Anabaena* 7120 *psbA* and *woxA* have only one transcriptional start [17, D. Borthakur, personal communication]. The *woxA* gene encodes the 33 kDa manganese-binding protein of the water-splitting complex associated with PSII.

Close examination of the region preceding the *Anabaena* 7120 *psbB* −238 start shows sequences that can be found upstream of the *Synechocystis* 6803 *psbB* gene at positions −169 to −203 (Fig. 6b). These sequences may be important in regulating transcription. There are additional sequences after the −238 start, at positions −235 to −220, that are also found upstream of the *Synechocystis* 6803 *psbB* gene at positions −161 to −146. The *Synechocystis* 6803 *psbB* transcripts have not been mapped and so it is not known where they start or whether there are also two promoters for the gene.

Sequences in front of the other *Anabaena* 7120 *psbB* start at −302 can be found upstream of other *Anabaena* 7120 genes. The sequences TTGTAA and ATAAA can be found near the proposed transcript starts for the phycocyanin operon and *atpB/E* genes as well as the *psbB* gene (Fig. 6a). The significance of these shared sequences is unknown since only the region upstream of the *psbB* −302 start has been shown *in vitro* to contain a promoter [48].

A comparison of all of the cloned *Anabaena* 7120 vegetative cell genes reveals a poor consensus concerning transcription initiation. Transcripts can start as close as 67 bp from the start codon for *psbA* and as far as 414 bp away for the *rbcLS* genes [17, 43]. The *Anabaena* 7120 genes cloned so far do not have sufficiently conserved sequences near their proposed transcriptional start sites to identify a promoter consensus

sequence readily [48]. Further experiments involving deletions are underway to better define the *psbB* gene promoter.

The 3′ end of the *psbB* transcripts were also mapped and found to follow a 21 bp inverted repeat that begins 239 bases after the stop codon. It seems likely that the repeat can form a stable secondary structure in single-stranded DNA since it was not possible to sequence through this region using the Sanger sequencing method with either *E. coli* DNA polymerase I (Klenow) or modified T7 DNA polymerase (Sequenase). The region was finally sequenced using reverse transcriptase, an enzyme known to be capable of polymerization through regions containing stable secondary structures [51]. The 21 bp inverted repeat at the 3′ end of the *psbB* gene transcripts may be similar to those stem-loop structures following several bacterial and chloroplast genes that act as stabilizing elements [42, 52].

Another stem might be formed by the repeats beginning 6 bp after the stop codon. This structure would be composed of 7 bp repeats. Short repeats like these are also found between the *Anabaena* 7120 *atpB/E* genes [15], *cpcEF* genes [4], *nifB* and *fdxN* genes, *nifS* and *nifU* genes (M.E. Mulligan, personal communication), and following the *petF* gene [2] and the *nifK* gene [35]. Since there is no evidence for transcript termini just after the repeats following the *psbB* stop codon, these repeats do not form a processing site. These repeats and the ones after the *petF* gene may, however, slow down the RNA polymerase and allow termination further downstream (after the 21 bp inverted repeat for *psbB*).

The question of whether *psbB* transcripts can be found in heterocysts still exists since it is difficult to isolate RNA from heterocysts. The procedure used to lyse vegetative cells with lysozyme results in the isolation of somewhat degraded RNA from the remaining heterocysts. We expect that when intact RNA is isolated from heterocysts, the transcripts for *psbB* will be absent as has recently been shown for *rbcLS* [3, 23].

It was possible to examine the state of *psbB* transcripts during the induction period. The level of transcripts for the *psbB* gene is constant during the induction period as it is for the transcripts of the *psbA*, *rbcLS* and *glnA* genes. In early experiments to study the transcription of *rbcLS* during induction, it was found that the transcripts disappear [28]. This difference may be due to the method used to induce nitrogen fixation. In the earlier experiments, an anaerobic induction was done with argon and DCMU. This creates an artificial condition that allows transcription of *nifHDK* in semidifferentiated cells. If oxygen is added, nitrogen fixation stops and the *nifHDK* message disappears. In the experiments reported here, the induction was done aerobically by transferring the culture from N+ to N− media, causing heterocysts to differentiate fully. Since the second induction represents a more natural condition, these results may be a better reflection of the true state of transcription during heterocyst differentiation. The only changes in transcription seen so far are those that involve the expression of genes that are turned on at different times during the induction like the *nif* genes and more recently *fdxH*, which encodes a heterocyst-specific ferredoxin [9]. When interpreting these results, one must consider that total RNA is being isolated and that approximately 10% of the cells in a differentiating culture will become heterocysts. Any decreases in the transcription of genes in heterocysts may not be seen because of continued transcription of the same genes in the vegetative cells.

Acknowledgements

We thank R.G. Herrmann for the spinach *psbB* probe, W.R. Belknap and S.A. Nierzwicki-Bauer for advice to J.D.L. throughout this work, J. Golden for the RNA samples, M.E. Mulligan for discussion concerning repeated sequences, and W.J. Buikema and F.L.H. Wolfs for computer assistance. The work was supported by grant GM21823 from the N.I.H.. J.D.L. was a predoctoral trainee in Molecular and Cell Biology (GM07183).

References

1. Adams DG, Carr NG: Developmental biology of heterocyst and akinete formation in cyanobacteria. Crit Rev Microbiol 9: 45–100 (1981).
2. Alam J, Whitaker RA, Krogmann DW, Curtis SE: Isolation and sequence of the gene for Ferredoxin I from the cyanobacterium *Anabaena* sp. strain PCC 7120. J Bact 168: 1265–1271 (1986).
3. Almon H, Böhme H: Components and activity of the photosynthetic electron transport system of intact heterocysts isolated from the blue-green alga *Nostoc muscorum*. Biochim Biophys Acta 592: 113–120 (1980).
4. Belknap WR, Haselkorn R: Cloning and light regulation of expression of the phycocyanin operon of the cyanobacterium *Anabaena*. EMBO J 6: 871–884 (1987).
5. Benton WD, Davis RW: Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science 196: 180–181 (1977).
6. Bergman B, Lindblad P, Pettersson A, Renström E, Tiberg E: Immuno-gold localization of glutamine synthetase in a nitrogen fixing cyanobacterium (*Anabaena cylindrica*). Planta 166: 329–334 (1985).
7. Bergman B, Lindblad P, Rai AN: Nitrogenase in free-living and symbiotic cyanobacteria: immunoelectron microscopic localization. FEMS Microbiol Lett 35: 75–78 (1986).
8. Bergman B, Rai AN, Söderbäck E: *Nostoc* of the *Anthoceros* and *Gunnera* symbioses: Immunoelectron microscopic localization of key proteins. Proceedings Sixth International Symposium on Photosynthetic Prokaryotes: 50 (1988).
9. Böhme H, Haselkorn R: Molecular cloning and nucleotide sequence analysis of the gene coding for heterocyst ferredoxin from the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol Gen Genet 214: 278–285 (1988).
10. Bolivar F, Backman K: Plasmids of *Escherichia coli* as cloning vectors. In: Wu R (ed) Methods in Enzymology, vol 68, pp. 245–267. Academic Press, New York (1979).
11. Braun-Howland EB, Nierzwicki-Bauer SA: Localization of the 32 kd protein in an *Azolla* endosymbiont. Proceedings Sixth International Symposium on Photosynthetic Prokaryotes: 122 (1988).
12. Casadaban M, Cohen SN: Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J Mol Biol 138: 179–207 (1980).
13. Chan TM, Markley JL: Nuclear magnetic resonance studies of two-iron-sulfur ferredoxins. 1. Properties of the histidine residues. Biochemistry 22: 5982–5987 (1983).
14. Codd GA, Stewart WDP: Ribulose-1,5-diphosphate carboxylase in heterocysts and vegetative cells of *Anabaena cylindrica*. FEMS Microbiol Lett 2: 247–249 (1977).
15. Curtis SE: Genes encoding the beta and epsilon subunits of the proton-translocating ATPase from *Anabaena* sp. strain PCC 7120. J Bact 169: 80–86 (1987).
16. Curtis SE, Haselkorn R: Isolation and sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from the cyanobacterium *Anabaena* 7120. Proc Natl Acad Sci USA 80: 1835–1839 (1983).
17. Curtis SE, Haselkorn R: Isolation, sequence and expression of two members of the 32 kd thylakoid membrane protein gene family from the cyanobacterium *Anabaena* 7120. Plant Mol Biol 3: 249–258 (1984).
18. Deisenhofer J, Epp O, Miki K, Huber R, Michel H: Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution. Nature 318: 618–624 (1985).
19. Deisenhofer J, Epp O, Miki K, Huber R, Michel H: X-ray structure analysis of a membrane protein complex. J Mol Biol 180: 385–398 (1984).
20. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395 (1984).
21. Fleming H, Haselkorn R: Differentiation in *Nostoc muscorum*: nitrogenase is synthesized in heterocysts. Proc Natl Acad Sci USA 70: 2727–2731 (1973).
22. Fleming H, Haselkorn R: The program of protein synthesis during heterocyst differentiation in nitrogen fixing blue-green algae. Cell 3: 159–170 (1974).
23. Golden JM, Carrasco CD, Mulligan ME, Schneider GJ, Haselkorn R: Deletion of a 55 kb DNA element from the chromosome during heterocyst differentiation of *Anabaena* sp. strain PCC 7120. J Bact 170: 5034–5041 (1988).
24. Golden JM, Robinson SJ, Haselkorn R: Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. Nature 314: 419–423 (1985).
25. Golden SS, Brusslan J, Haselkorn R: Expression of a family of *psbA* genes encoding a photosystem II polypeptide in the cyanobacterium *Anacystis nidulans* R2. EMBO J 5: 2789–2798 (1986).
26. Golden SS, Brusslan J, Haselkorn R: Genetic engineering of the cyanobacterial chromosome. In: Wu R, Grossman L (eds) Methods in Enzymology, vol 153, pp. 215–231. Academic Press, San Diego (1987).
27. Haselkorn R: Heterocysts. Ann Rev Plant Physiol 29: 319–344 (1978).
28. Haselkorn R, Rice D, Curtis SE, Robinson SJ: Organization and transcription of genes important in *Anabaena* heterocyst differentiation. Ann Microbiol 134B: 181–193 (1983).
29. Kratz WA, Myers J: Nutrition and growth of several blue-green algae. Am J Bot 42: 282–287 (1955).
30. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–132 (1982).
31. Lammers PJ, Haselkorn R: Sequence of the *nifD* gene coding for the α subunit of dinitrogenase from the cyanobacterium *Anabaena*. Proc Natl Acad Sci USA 80: 4723–4727 (1983).
32. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning:

A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
33. Marck C: 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res 16: 1829–1836 (1988).
34. Maxam AM, Gilbert W: Sequencing end-labeled DNA with base-specific chemical cleavages. In: Grossman L, Moldave K (eds) Methods in Enzymology, vol 65, pp. 499–559. Academic Press, New York (1980).
35. Mazur BJ, Chui CF: Sequence of the gene coding for the β-subunit of dinitrogenase from the blue-green alga *Anabaena*. Proc Natl Acad Sci USA 79: 6782–6786 (1982).
36. Mazur BJ, Rice D, Haselkorn R: Identification of blue-green algal nitrogen fixation genes using heterologous DNA hybridization probes. Proc Natl Acad Sci USA 77: 186–190 (1980).
37. Mevarech M, Rice D, Haselkorn R: Nucleotide sequence of a cyanobacterial *nifH* gene coding for nitrogenase reductase. Proc Natl Acad Sci USA 77: 6476–6480 (1980).
38. Morris J, Herrmann RG: Nucleotide sequence of the gene for the P_{680} chlorophyll *a* apoprotein of the photosystem II reaction center from spinach. Nucleic Acids Res 12: 2837–2850 (1984).
39. Mulligan ME, Buikema WJ, Haselkorn R: Bacterial-type ferredoxin genes in the nitrogen-fixation regions of the cyanobacterium *Anabaena* sp. strain PCC 7120 and *Rhizobium meliloti*. J Bact 170: 4406–4410 (1988).
40. Nakatani HY, Ke B, Dolan E, Arntzen CJ: Identity of the photosystem II reaction center polypeptide. Biochim Biophys Acta 765: 347–352 (1984).
41. Nanba O, Satoh K: Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome b-559. Proc Natl Acad Sci USA 84: 109–112 (1987).
42. Newbury SF, Smith NH, Higgins CF: Differential mRNA stability controls relative gene expression within a polycistronic operon. Cell 51: 1131–1143 (1987).
43. Nierzwicki-Bauer SA, Curtis SE, Haselkorn R: Cotranscription of genes encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in the cyanobacterium *Anabaena* 7120. Proc Natl Acad Sci USA 81: 5961–5965 (1984).
44. Norrander J, Kempe T, Messing J: Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene 26: 101–106 (1983).
45. Orr J, Haselkorn R: Regulation of glutamine synthetase activity and synthesis in free-living and symbiotic *Anabaena* spp. J Bact 152: 626–635 (1982).
46. Rice D, Mazur BJ, Haselkorn R: Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120. J Biol Chem 257: 13157–13163 (1982).
47. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
48. Schneider GJ, Lang JD, Haselkorn R: *Anabaena* promoters defined by *in vitro* transcription. Manuscript in preparation.
49. Schrautemeier B, Böhme H: A distinct ferredoxin for nitrogen fixation isolated from heterocysts of the cyanobacterium *Anabaena variabilis*. FEBS Lett 184: 304–308 (1985).
50. Shine J, Dalgarno L: The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA 71: 1342–1346 (1974).
51. Smith AJH: DNA sequence analysis by primed synthesis. In: Grossman L, Moldave K (eds) Methods in Enzymology, vol 65, pp. 560–580. Academic Press, New York (1980).
52. Stern DB, Gruissem W: Control of plastid gene expression: 3′ inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. Cell 51: 1145–1157 (1987).
53. Taber S, Richardson CC: DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc Natl Acad Sci USA 84: 4767–4771 (1987).
54. Thomas J: Absence of pigments of photosystem II of photosynthesis in heterocysts of a blue-green alga. Nature 258: 715–716 (1970).
55. Thomas PS: Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. In: Wu R, Grossman L, Moldave K (eds) Methods in Enzymology, vol 100, pp. 255–266. Academic Press, New York (1983).
56. Tronrud DE, Schmid MF, Matthews BW: Structure and X-ray amino acid sequence of a bacteriochlorophyll *a* protein from *Prosthecochloris aestuarii* refined at 1.9 Å resolution. J Mol Biol 188: 443–454 (1986).
57. Tumer NE, Robinson SJ, Haselkorn R: Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. Nature 306: 337–341 (1983).
58. Vermaas WFJ, Ikeuchi M, Inoue Y: Protein composition of the photosystem II core complex in genetically engineered mutants of the cyanobacterium *Synechocystis* sp. PCC 6803. Photosynthetic Res 17: 97–113 (1988).
59. Vermaas WFJ, Williams JGK, Arntzen CJ: Sequencing and modification of *psbB*, the gene encoding the CP-47 protein of Photosystem II, in the cyanobacterium *Synechocystis* 6803. Plant Mol Biol 8: 317–326 (1987).
60. Winkenbach F, Wolk CP: Activities of enzymes of the oxidative and the reductive pentose phosphate pathways in heterocysts of a blue-green alga. Plant Physiol 52: 480–483 (1973).
61. Wood NB, Haselkorn R: Control of phycobiliprotein proteolysis and heterocyst differentiation in *Anabaena*. J Bact 141: 1375–1385 (1980).
62. Yamagishi A, Katoh S: A photoactive photosystem-II reaction center complex lacking a chlorophyll-binding 40