

# Stringent promoter recognition and autoregulation by the group 3 $\sigma$ -factor SigF in the cyanobacterium *Synechocystis* sp. strain PCC 6803

Munehiko Asayama\* and Sousuke Imamura

Laboratory of Molecular Genetics, School of Agriculture, Ibaraki University, 3-21-1 Ami, Inashiki, Ibaraki 300-0393, Japan

Received April 8, 2008; Revised June 17, 2008; Accepted June 30, 2008

## ABSTRACT

The cyanobacterium *Synechocystis* sp. strain PCC 6803 possesses nine species of the sigma ( $\sigma$ )-factor gene for RNA polymerase (RNAP). Here, we identify and characterize the novel-type promoter recognized by a group 3  $\sigma$ -factor, SigF. SigF autoregulates its own transcription and recognizes the promoter of *pilA1* that acts in pilus formation and motility in PCC 6803. The *pilA1* promoter (PpilA1-54) was recognized only by SigF and not by other  $\sigma$ -factors in PCC 6803. No PpilA1-54 activity was observed in *Escherichia coli* cells that possess RpoF ( $\sigma^{28}$ ) for flagellin and motility. Studies of *in vitro* transcription for PpilA1-54 identified the region from –39 to –7 including an AG-rich stretch and a core promoter with TAGGC (–32 region) and GGTA (–12 region) as important for transcription. We also confirmed the unique PpilA1-54 architecture and further identified two novel promoters, recognized by SigF, for genes encoding periplasmic and phytochrome-like phototaxis proteins. These results and a phylogenetic analysis suggest that the PCC 6803 SigF is distinct from the *E. coli* RpoF or RpoD ( $\sigma^{70}$ ) type and constitutes a novel eubacterial group 3  $\sigma$ -factor. We discuss a model case of stringent promoter recognition by SigF. Promoter types of PCC 6803 genes are also summarized.

## INTRODUCTION

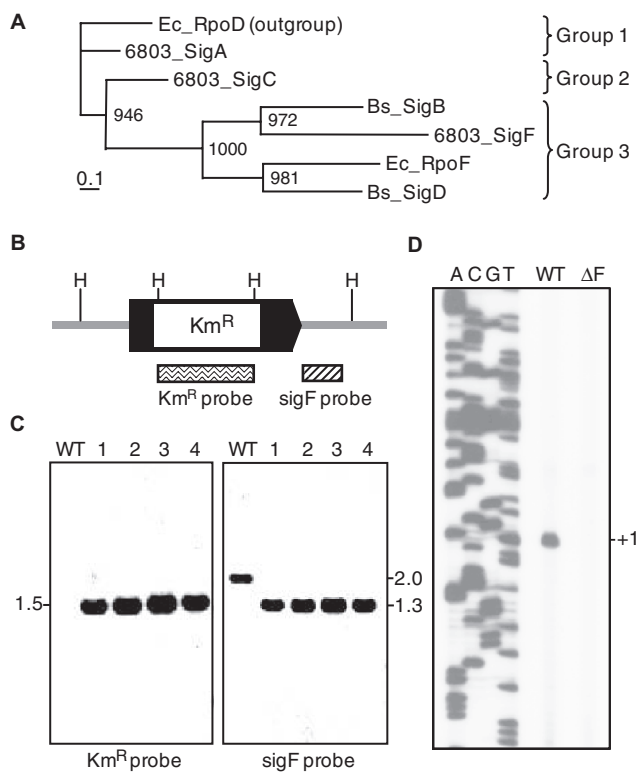
The RNA polymerase (RNAP) holoenzyme of eubacteria consists of a core enzyme and sigma ( $\sigma$ )-factor (1,2). The core enzyme confers the capability of RNA synthesis and the  $\sigma$ -factor is required for the initiation of transcription from a promoter sequence. The  $\sigma$ -factors can be divided into two families,  $\sigma^{70}$  and  $\sigma^N$  types, in *Escherichia coli* (3,4). The  $\sigma^{70}$  family can be structurally and functionally

subdivided into three groups (4). Group 1 comprises a principal  $\sigma$ -factor that is essential for cell viability. Group 2  $\sigma$ -factors are similar to group 1 in molecular structure, but are nonessential for cell viability. Group 3  $\sigma$ -factors are an alternative type, structurally different from group 1 and group 2 factors, and are involved in the transcription of regulons for survival under stress.

Cyanobacteria are Gram-negative prokaryotes that perform oxygenic photosynthesis like plants. It is generally accepted that an ancestor of cyanobacteria gave rise to plant chloroplasts through a unique endosymbiotic event, thereby conferring the ability for photosynthesis to algae and plants. Complete nucleotide sequences of several cyanobacterial genomes have been decoded, indicating that no  $\sigma^N$  types exist in cyanobacteria. The non-nitrogen-fixing unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803 possesses nine species of  $\sigma$ -factor assigned to group 1 (SigA), group 2 (SigB, SigC, SigD and SigE) and group 3 (SigF, SigG, SigH and SigI) (5–8). The functions of the group 2  $\sigma$ -factors have been revealed recently. For example, SigD/SigB are light-/dark-induced  $\sigma$ -factors and SigB was also identified as a heat shock-responsive  $\sigma$ -factor (6,9–11). SigC and SigB contribute to the transcription of nitrogen metabolism-related genes, depending on the phase of cell growth (12,13). SigE is required for survival under nitrogen stress and positive regulation of sugar catabolic pathways (13–16). Light-induced transcription also involves a rhythmic expression, in which the periodic peak of SigE exhibits a 24-h interval according to the upcoming night (17). The diversity of function in group 2 heterogeneous  $\sigma$ -factors is characteristic of cyanobacteria. Furthermore, group 1 and 2  $\sigma$ -factors coexist and cooperate in interfering with the network (18) of transcription from promoters whose potential sequences are of the *E. coli*  $\sigma^{70}$  type.

Promoter sequences recognized by group 3  $\sigma$ -factors in photosynthesizing eubacteria have not been studied. SigF (Slr1564, Figure 1A) is the only  $\sigma$ -factor whose function has been identified as concerned with pilus formation and motility among PCC 6803 group 3  $\sigma$ -factors

\*To whom correspondence should be addressed. Tel./Fax: +81 29 888 8651; Email: asam@mx.ibaraki.ac.jp.



**Figure 1.** SigF type and *pilA1* expression in the *sigF* knockout strain. (A) Phylogenetic analysis of SigF. The numbers at each node indicate the percent of trees supporting the specific branching pattern in the bootstrap (1000 times) analysis. The bar indicates the distances corresponding to 20 changes per 100 amino acid positions; 6803, *Synechocystis* PCC 6803; Bs, *Bacillus subtilis*; Ec, *Escherichia coli*. (B) Gene disruption of *sigF* (filled arrow) was performed by insertion of the kanamycin-resistance gene ( $Km^R$ , white box). Each probe for genomic Southern analysis is shown. H, *Hind*III. (C) Total DNA (5  $\mu$ g) was prepared from PCC 6803 WT or the *sigF* knockout strains ( $\Delta F$ , numbers 1–4), and digested with *Hind*III, and genomic Southern analysis was carried out with the probes. The signal size is shown as kilo-bases on the right or left. (D) Total RNA (10  $\mu$ g) was prepared from WT or  $\Delta F$  strains, and then 5'-end mapping of the *pilA1* transcript was conducted by primer extension with the *pilA1*-R primer (5'-TTG GAGAGTTGAGAGAGGAGTTT-3'). The sequencing ladder was also synthesized with plasmid DNA, pPILA1000\_puc carrying 1000-bp nucleotides of the *pilA1* upstream region, and the same primer. The position of the transcription start point is indicated as +1 (PpilA1-54), referring to -54 from an initiation codon.

(Figure 1A) (7). Therefore, we tried to identify the core promoter sequence recognized by SigF as a model case in which the  $\sigma$ -factor autoregulates its own transcription. Based on previous results and the findings of this study, we also tried to classify PCC 6803 promoter types.

## MATERIALS AND METHODS

### Strains and media

*Synechocystis* sp. strain PCC 6803 was obtained from Kazusa DNA Research Institute. The cells were grown on BG11 plates or liquid medium, supplemented with kanamycin sulfate (Km, 15  $\mu$ g/ml) or chloramphenicol (Cm, 25  $\mu$ g/ml) as required under continuous white-light illumination (35  $\mu$ E m<sup>-2</sup> s<sup>-1</sup>). *Escherichia coli* strain

DH5 $\alpha$ <sup>MCR</sup>, which is used routinely for recombinant DNA manipulation and  $\beta$ -galactosidase assays (19,20), was grown on LB plates or liquid medium containing the appropriate antibiotic, ampicillin (Ap, 75  $\mu$ g/ml) or spectinomycin sulfate (Sp, 40  $\mu$ g/ml), if necessary.

### Plasmids

Upstream regions from -946 to +14, -346 to +14, -46 to +14 and +4 to +14 (+1 as the transcription start point) of PCC 6803 *pilA1* were amplified by PCR and cloned into the *Sma*I-*Bgl*II site of pUC119B (20–22), to create pPILA1000\_puc, pPILA400\_puc, pPILA76\_puc and pPILA20\_puc, respectively. The *pilA1* insert was in the reverse orientation to *lacZ* in each pPILA\_puc vector for the *in vitro* transcription analysis. Each insert of the *Sma*I-*Bgl*II fragment was isolated from these plasmids and inserted into the *Sma*I-*Bgl*II site of pAM990 (23), to obtain pPILA1000\_pam, pPILA400\_pam, pPILA76\_pam and pPILA20\_pam, respectively. The *pilA1* insert had the same orientation as *lacZ* in each pPILA\_pam (promoter-probe) vector for the  $\beta$ -galactosidase assay. Mutated versions of the *pilA1* promoter were constructed for a 7-bp (or 3-bp) scanning analysis as follows. A set of *pilA1m7*- (or *pilA1m3*-)Fw (76-mer, 5'-accccgggTGAGCA—TC CGACagatcttc-3': *Sma*I and *Bgl*II sites underlined, upper case is *pilA1* upstream sequence) and *pilA1m7*- (or *pilA1m3*-)Rv (76-mer, 5'-gaagatctGTCGGA—TGC TCaccgggggt-3') oligonucleotides were annealed, digested with *Sma*I and *Bgl*II, and then inserted into the *Sma*I-*Bgl*II site of pUC119B to yield pPILA76mt7-No'X' puc (or pPILA76mt3-No'X' puc), in which the positions of mutagenized sequences are shown in Figure 4A. The *Sma*I-*Bgl*II fragment was isolated from pPILA76mt7-No5\_puc and also inserted into the *Sma*I-*Bgl*II site of pAM990 to make pPILA76mt7-No5\_pam, shown in Figure 2D. The construction of pPILA61\_puc and pPILA61mt21 or mt22\_puc was done using the procedure described above for pPILA76\_puc, with *pilA1*-FwS61 (or *pilA1m*-Fw21/-Fw22, 61-mer; 5'-accccgggAGGGTG—CTAGAAagatcttc-3') and *pilA1*-RvS61 (or *pilA1m*-Rv21/-Rv22, 61-mer; 5'-gaagatctTTCTAG—CA CCCTcccgggggt-3'), respectively. The positions of mutagenized sequences in pPILA61mt21 and mt22 are shown in Figure 5A. The nucleotide sequences of inserts in the respective plasmids mentioned above were verified. pSIGF carrying the PCC 6803 *sigF*-promoter region and pAG500 containing the cyanobacterium *Microcystis aeruginosa* K-81 *psbA2* promoter were described previously (6,19,24).

### Proteins

All recombinant  $\sigma$ -factors of PCC 6803 were purified by a method described previously (6,10,13). Purification of the reconstituted PCC 6803 RNAP core enzyme with each recombinant subunit was performed as reported (10,13).

### Gene disruption and conjugative transformation

The disruption of the respective  $\sigma$ -factor genes, *sigB* to *sigI*, in strain PCC 6803 was described previously (6,18). The *sigF* knockout-strain ( $Km^R$ ) as a recipient was

subjected to conjugative transformation with helper *E. coli* cells carrying R751 (25) and DH5 $\alpha$ <sup>MCR</sup> harboring pVZ321SIGF-COMP (Cm<sup>R</sup>), which contains a region from -500 to +777 of *sigF* (+1 as a start codon of *sigF*), derived from pVZ321 (26). Km- and Cm-resistant PCC 6803 cells (Km<sup>R</sup>, Cm<sup>R</sup>) were selected and used for the complementation analysis.

### RNA isolation and primer extension analysis

These analyses were performed as described previously (21,22,27,28).

### In vitro transcription analysis

RNA was synthesized *in vitro* by multiple-round run-off transcription with the RNAP core enzyme (1 pmol),  $\sigma$  subunit (15 pmol) and template DNA (3  $\mu$ g), as reported previously (10–13,20,29). RNA products were precipitated with 2-propanol and dissolved in 10  $\mu$ l of RNase-free water. The samples were subjected to primer extension as above.

### Phylogenetic analysis

A phylogenetic tree was constructed by the method described previously (6). Total amino acid sequences of respective  $\sigma$ -factors were used as queries.

## RESULTS

### Phylogenetic analysis for SigF

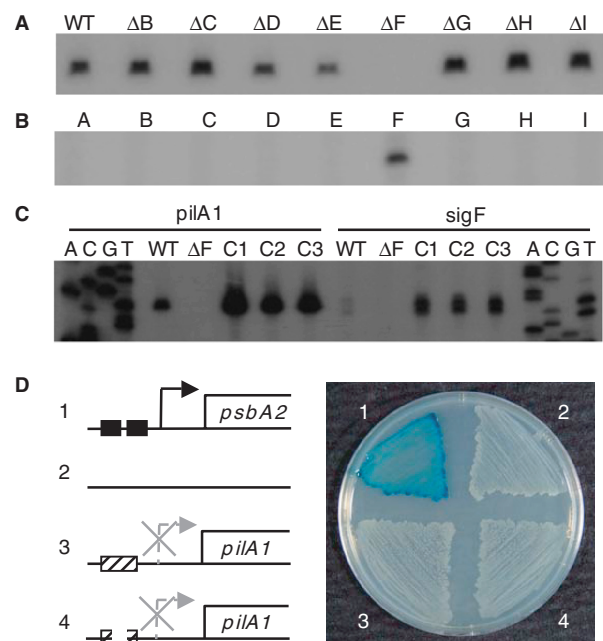
The molecular structure of PCC 6803 SigF was analyzed by analogy of amino acid sequences and the result is shown in Figure 1A. The SigF protein is assigned as a group 3  $\sigma$ -factor but might not be the ECF type in PCC 6803 (6,9,18,30). The analysis also shows that PCC 6803 SigF is in the same node as *Bacillus subtilis* SigB as a general stress-responsive  $\sigma$ -factor but not in the node that includes *B. subtilis* SigD and *E. coli* SigF as  $\sigma$ -factors for chemotaxis and flagellin.

### 5'-End mapping of the *pilA1* transcript

A previous study using northern blotting indicated that *pilA1* (*sll1694*) was not transcribed in the SigF knockout mutant (7), indicating that SigF contributes to the transcription process in PCC 6803. In a further analysis, we characterized the *pilA1* transcripts in SigF knockout mutants. A kanamycin-resistance gene (Km<sup>R</sup>) was inserted into the *sigF* gene in the PCC 6803 genome (Figure 1B) and knockout mutants were confirmed by genomic Southern analysis (Figure 1C). The 5'-end mapping of the *pilA1* transcript by primer extension revealed a single transcription start point (+1), from promoter PpilA1-54, in wild-type cells (WT), but no signal in the knockout mutant ( $\Delta F$ ) (Figure 1D). This supports a previous finding that SigF is required for the transcription of *pilA1* (7).

### Stringent promoter recognition by SigF

We tested for PpilA1-54 activity among mutants in which each  $\sigma$ -factor gene was knocked out (Figure 2A).



**Figure 2.** Stringent promoter recognition of SigF. (A) Primer extension was performed with the *pilA1*-R primer and total RNA (10  $\mu$ g) prepared from 6803 WT or the respective  $\sigma$ -factor disrupted gene.  $\Delta B$ ,  $\Delta sigB$ ;  $\Delta C$ ,  $\Delta sigC$ ;  $\Delta D$ ,  $\Delta sigD$ ;  $\Delta E$ ,  $\Delta sigE$ ;  $\Delta F$ ,  $\Delta sigF$ ;  $\Delta G$ ,  $\Delta sigG$ ;  $\Delta H$ ,  $\Delta sigH$  and  $\Delta I$ ,  $\Delta sigI$ . (B) *In vitro* transcription was performed with recombinant 6803 RNAP reconstituted with the core enzyme and each  $\sigma$ -factor (A, SigA; B, SigB; C, SigC; D, SigD; E, SigE; F, SigF; G, SigG; H, SigH and I, SigI). Transcripts synthesized from the template plasmid DNA, pPILA76\_puc carrying 76 bp of the *pilA1* upstream region, were analyzed by primer extension with the RV primer (5'-CAGGAAACAGCTATGAC-3'). (C) Primer extension of the *pilA1* or *sigF* gene was performed with the *pilA1*-R or *sigF*-R1 primer (6) and total RNA (10  $\mu$ g) prepared from 6803 WT,  $\Delta F$  or  $\Delta F$  complementant 6803 RNAP (C1, C2 and C3), respectively. Signals correspond to transcription start points (+1) of PpilA1-54 or PsigF-21/-22. Sequencing ladders were also synthesized with plasmid DNA, pPILA1000\_pam or pSIGF, and the same *pilA1*-R or *sigF*-R1 primer, respectively. The X-ray film was subjected to autoradiography for 1 day. (D) The *pilA* promoter activity in *E. coli* DH5 $\alpha$ <sup>MCR</sup> cells harboring pAG500<1>, pAM990<2>, pPILA76\_pam <3> or pPILA76m7-No5\_pam <4>, respectively, was observed on the LB plate supplemented with ampicillin (75  $\mu$ g/ml), spectinomycin (40  $\mu$ g/ml) and X-Gal (0.02%, w/v). These cells were incubated for 1 day at 37°C and then for 3 days at 4°C. Representative promoter (box)/transcription (arrow) and images of the cells are shown on the left and right, respectively.

PpilA1-54 activity was lost in the SigF mutant but was observed in the others. This shows that of all the 6803  $\sigma$ -factors, only SigF is essential for PpilA1-54 transcription in the cell. In addition, the signal intensity was slightly reduced in the  $\Delta D$  and  $\Delta E$  strains, suggesting an indirect contribution to PpilA1-54 activity in the cell. Which  $\sigma$ -factors can directly recognize the *pilA1* promoter? To answer this question, an *in vitro* transcription analysis was performed using RNAPs, reconstituted with the PCC 6803 core enzyme and each  $\sigma$ -factor (Figure 2B). A specific signal for PpilA1-54 was confirmed when RNAP-SigF was used. On the other hand, no signal was observed when RNAPs were examined with other  $\sigma$ -factors. These results show that only RNAP-SigF can directly recognize PpilA1-54 *in vitro*. RNAP-SigF is therefore

sufficient for PpilA1-54 basal transcription and other transcription factors are unnecessary.

To confirm SigF's function in PCC 6803 cells, we conducted a complementation test (Figure 2C, left). The *sigF* gene, cloned in a vector, was conjugatively transferred into the  $\Delta F$  strain. Signal for PpilA1-54 was observed in WT but not in  $\Delta F$ . Strong signals were again observed in  $\Delta F$  strains compensated with *sigF* (C1, C2, and C3), indicating a specific function of SigF in the cells. Of note, the strong signal intensity might be dependent on multi-copies of *sigF* and its high level of expression in C1, C2 and C3 cells.

### Autoregulation of SigF expression

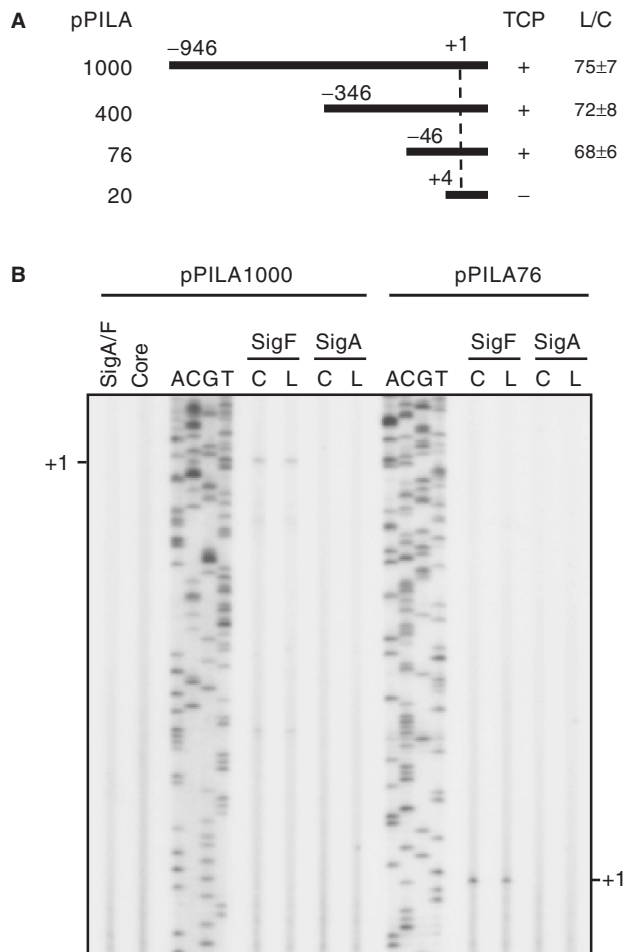
It has been reported that large amounts of the *sigF* transcript were not detected in the SigF mutant by QRT-PCR (18). To identify another promoter recognized by SigF and to elucidate whether autoregulation of SigF expression exists or not in PCC 6803, we experimented further (Figure 2C, right). The transcription start points of P*sigF*-22/-21, which were previously reported (6), were observed in WT but not in  $\Delta F$ . P*sigF*-22/-21 activities were confirmed in C1, C2 and C3 cells. We also confirmed that P*sigF*-22/-21 transcription was specifically driven *in vitro* by RNAP-SigF (data not shown), clearly showing that *sigF* expression is autoregulated.

### *E. coli* $\sigma$ -factors cannot recognize the *pilA1* promoter

The Gram-negative bacterium *E. coli* has a  $\sigma$ -factor, RpoF ( $\sigma^F$ ,  $\sigma^{28}$ ), which participates in flagellar formation and motility. Cyanobacteria are also Gram-negative organisms and have a group of  $\sigma$ -factors that are involved in the transcription of genes required for motility. For example, PCC 6803 SigF can recognize the promoter of the *pilA1* gene that plays a role at the cellular surface. Therefore, to characterize PpilA1-54 expression in *E. coli*, we constructed several fusions of *pilA1::lacZ*, as derivatives from a promoter-probe vector, pAM990 (Figure 2D, left). They were individually introduced into *E. coli* cells and  $\beta$ -galactosidase activity, according to *pilA1* promoter activity, was observed on the LB plate (Figure 2D, right). When we used pAG500 (K-81 *psbA2::lacZ*) (19,20,24) as a positive control,  $\beta$ -galactosidase activity was observed. Meanwhile, no activity was observed from pAM990 (vector only) and pPILA76m7-No5\_pam (mutagenized 6803 *pilA1::lacZ*) as negative controls. PpilA1-54 activity could not be detected from pPILA76\_pam (6803 *pilA1::lacZ*). Of note, heterologous RNAP, reconstituted with the *E. coli* core enzyme and PCC 6803 SigF, can recognize PpilA1-54 *in vitro* (shown later in Figure 5C). These results indicate that *E. coli*  $\sigma$ -factors cannot recognize the *pilA1* promoter in *E. coli* cells.

### An essential region for PpilA1-54 transcription

To evaluate the promoter, PpilA1-54, *in vitro* transcription analyses by primer extension were performed with deletion constructs (Figure 3A). When RNAP-SigF and each plasmid construct were mixed and reacted *in vitro*, apparent specific PpilA1-54 transcripts were detected in the



**Figure 3.** Minimal *pilA1* promoter region and efficient transcription by RNAP-SigF from supercoiled DNA templates. (A) *In vitro* transcription was performed with RNAP-SigF and each plasmid DNA template of pPILA1000\_puc, pPILA400\_puc, pPILA76\_puc and pPILA20\_puc. Synthesized transcripts were analyzed by primer extension with the RV primer. TCP: +, transcription; -, no transcription. The ratio (%) of the amount of transcripts, synthesized from the *Sma*I-linearized (L) or supercoiled (C) template DNAs of pPILA, is shown ( $n=3$ , means  $\pm$ SD). (B) The *pilA1* transcripts were synthesized as described in (A) and some of the results are presented. SigA/F, only a fraction of SigA and SigF without the RNAP core enzyme (Core). Sequencing ladders were also made with template DNAs and the RV primer. The position of the transcription start point is indicated as +1. Signal intensities referring to +1 were measured as described previously.

cases of pPILA1000 (*pilA1* region containing from -946 to +14, +1 as transcription start point, see the result in Figure 3B), pPILA400 (-346 to +14) and pPILA76 (-46 to +14, see the result in Figure 3B), respectively. However, no signal was detected when pPILA20 (+4 to +14) was used as the DNA template (data not shown, but see Figure 3A). These results indicate that the region from -46 to +14 is at least important for *in vitro* PpilA1-54 transcription.

### Template DNA form-dependent transcription by SigF

Generally, transcription depends upon template DNA forms. In cyanobacteria, supercoiled DNA templates of

the PCC 6803 *psbA2* promoter were more effectively transcribed than linear templates by RNAP-SigA (6). We therefore compared transcription efficiency using supercoiled (C) or linear (L) DNA templates (Figure 3A). The results obtained using pPILA1000 and 76 are shown in Figure 3B. Interestingly, 68–75% of transcripts were synthesized from linear versus supercoiled (L/C) templates (Figure 3A), showing that RNAP-SigF prefers supercoiled DNA templates. Of note, when RNAP-SigA was used as a control, transcripts from PpilA1-54 could be detected on neither supercoiled nor linear DNA templates (Figure 3B). This again supports stringent promoter recognition by SigF (Figure 2).

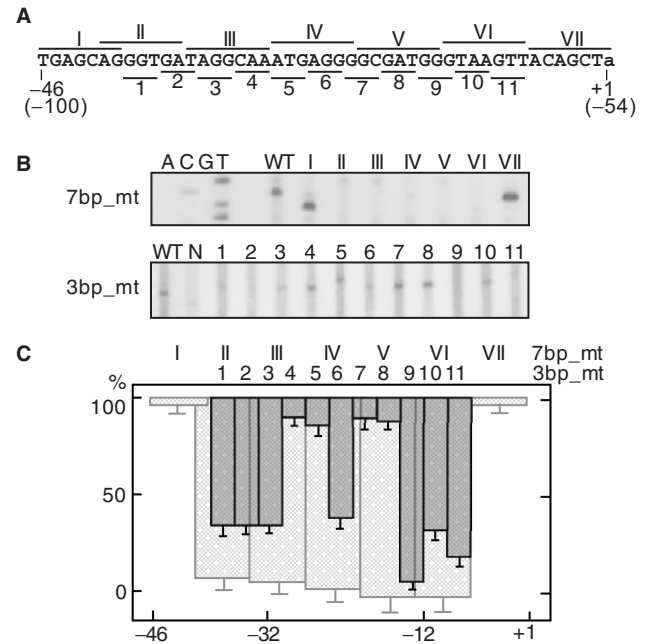
### Core sequence of PpilA1-54 for RNAP-SigF transcription

To identify the promoter sequence for PpilA1-54, we prepared mutant constructs for scanning analysis (Figure 4A) and performed *in vitro* transcription analyses (Figure 4B). In this assay, we used supercoiled pPILA76mt7-No'X' puc plasmids containing the PpilA1-54 region from -46 to +14, which possesses a mutagenized 7-bp sequence of transversion change (A↔T or C↔G, see the positions in Figure 4A top). For example, the sequence TGAGCAG was changed to ACTCGTC in pPILA76mt7-NoI puc. Therefore, transcripts of the same length might be synthesized *in vitro* from each derivative, and signal intensity according to the transcripts can be compared. No or very few transcripts were synthesized from mt7-NoII to NoVI, indicating that the region from -41 to -7 is important for PpilA1-54. The specific transcript but an inaccurate transcription start point was observed when mt7-NoI was used, suggesting that the sequence from -46 to -40 affects the selection of the transcription start site. In order to further understand the core promoter sequence, we consequently performed a scanning mutagenesis at a resolution of 3 bp with 11 different mutants of pPILA76mt3-No'X' puc, which contain the transversion change (A↔T or C↔G, see the positions in Figure 4A bottom). In this assay, no or very weak activity for transcription was detected when we used mt3-No1 to No3, No6 and No9 to No11. Of note, the specific transcript but inaccurate transcription start points of PpilA1-55 were also observed when mt3-No5 and mt3-No10 were used. These nucleotide substitutions in the spacer region or upstream from -32 might be affective for accurate transcription from PpilA1-54.

Results of the 7-bp and 3-bp scanning analyses are summarized in Figure 4C. Both regions of -32 and -12 might be the most significant for the transcription. Thus, we concluded tentatively that PpilA1-54 may comprise a core promoter sequence containing the -32 and -12 regions. The results also suggested that the region from -39 to -37 is an important *cis*-element for PpilA1-54 transcription, since the level of transcription was significantly lower when mt3-No1 was used.

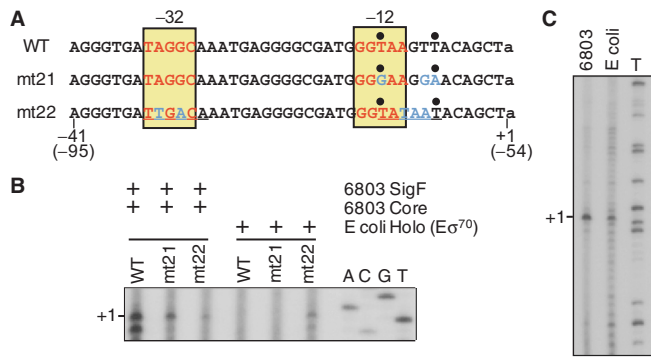
### Transcription on mutagenized PpilA1-54 by PCC 6803 or *E. coli* RNAP

The possible core sequence of PpilA1-54 is unique but the region also seems to have a set of sequences at -31



**Figure 4.** Identification of a core sequence of the *pilA1* promoter. (A) DNA templates for scanning mutagenesis of PpilA1-54 at a 7-bp (top) and 3-bp (bottom) resolution are shown. The supercoiled DNA templates of pPILA76\_puc (WT) or its derivatives (the positions of each mutation are numbered). (B) Primer extension was conducted with the RV primer and the transcripts synthesized *in vitro* with RNAP-SigF and DNA templates. N, vector pUC119B was used as a negative control. (C) Signal intensities from respective samples in (B) were measured and presented as relative values (WT as 100%) with error bars ( $n = 3$ , means  $\pm$  SD). The number of constructs and the position of the *pilA1* upstream region are indicated at the top and bottom, respectively.

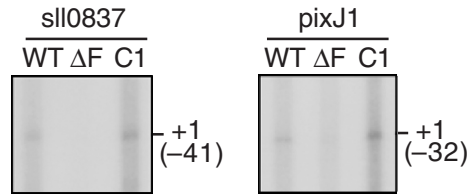
(TAGGCA) and -9 (TAAGTT) that exhibits a low degree of similarity to that of the cyanobacterial (and *E. coli*) consensus promoter as -35 (TTGACA) and -10 (TATAAT) hexamers. However, the set of pseudo sequences, overlapping partially in the PpilA1-54 region, does not function as the *E. coli*  $\sigma^{70}$ -type promoter because no transcription or expression was observed *in vitro* and *in vivo* by *E. coli*  $\sigma^{70}$ -type RNAPs (Figures 2 and 3). To further clarify this point, we inspected the distinct ability of promoter recognition by PCC 6803 or *E. coli* RNAP using mutagenized DNA templates of PpilA1-54 (Figure 5). The DNA template of pPILA61mt21\_puc (Figure 5A) has substituted nucleotides, guanine (G) or adenine (A) for thymine (T), at -12, -8 and -7 which might be preferred or allow transcription by RNAP-SigF but not by *E. coli*  $\sigma^{70}$ -type RNAPs. T-nucleotides at both ends of TATAAT (dots, TAAGTT in PpilA1-54) are essential and the substitution of T does not allow for transcription by the *E. coli*  $\sigma^{70}$ -type RNAP in cyanobacteria (21,22). On the other hand, the DNA template of pPILA61mt22\_puc is substituted in the promoter sequence for the *E. coli*  $\sigma^{70}$ -type with a 16-bp spacer. When RNAP reconstituted with the PCC 6803 core enzyme and SigF was used, specific transcription was driven *in vitro* from the WT of PpilA1-54 (-53) (Figure 5B). This transcription was allowed but was significantly low when mt21 was used. In the case of mt22,



**Figure 5.** Promoter recognition of mutagenized PpilA1-54 by RNAPs of PCC 6803 or *E. coli*. (A) DNA templates for *in vitro* transcription are shown. The supercoiled DNA templates of pPILA61\_puc (WT) or its derivatives (mt21 and 22, the nucleotides of each mutation are indicated as blue). The  $-32$  and  $-12$  regions are shown in boxes with red letters. The positions corresponding to thymine nucleotides at  $-12$  and  $-7$  are represented with dots. The promoter sequence of the *E. coli* consensus type (TTGACA and TATAAT) is underlined. (B) Primer extension was conducted with the RV primer and the transcripts synthesized *in vitro* with RNAPs (PCC 6803 core + PCC 6803 SigF, or *E. coli* holo) and DNA templates. ACGT is a sequencing ladder. (C) Primer extension was conducted with the RV primer and the transcripts synthesized *in vitro* with RNAPs (PCC 6803 core + PCC 6803 SigF, or *E. coli* core + PCC 6803 SigF) and the DNA template of pPILA76\_puc (WT) (Figures 3 and 4). T is a sequencing ladder.

the transcription by PCC 6803 RNAP-SigF was very limited. In the meantime, when an *E. coli* holoenzyme ( $E\sigma^{70}$ ) was used, specific transcription was observed in the case of mt22, while no transcription was observed in the cases of WT and mt21. These results indicate that the T-nucleotides in TAAGTT (partially as TATAAT) overlapping downstream from the  $-12$  core region of PpilA1-54 might be effective but not essential for the transcription by RNAP-SigF. Moreover, the T-nucleotide at  $-12$  may be important because the transcription is relatively weaker in the case of mt21 than WT. The results also indicated that the *E. coli* holoenzyme ( $E\sigma^{70}$ ) recognizes only the consensus type promoter sequence and not PpilA1-54. Therefore, we could confirm the unique architecture of PpilA1-54 comprising the core promoter sequences at  $-32$  and  $-12$ . Of note, the inaccurate transcription from PpilA1-53 observed in Figure 5B might be caused by using the short template DNA ( $-41$  to  $+4$ ) whose sequence lacks the region from  $-46$  to  $-42$  (see the result of pPILA76mt7-NoI\_puc shown in Figure 4).

We further reconfirmed promoter recognition by SigF using core enzymes of PCC 6803 or *E. coli*. The results are shown in Figure 5C. When RNAP reconstituted with the PCC 6803 core enzyme and SigF was used, specific transcription was observed from PpilA1-54. We also detected a major signal referring to the transcript from PpilA1-54 by heterologous RNAP reconstituted with the *E. coli* core enzyme and SigF. However, in this case, other minor signals were also observed. These results indicated that specific transcription depended on SigF and that the fidelity of transcription may be closely related to a property of the cyanobacterial RNAP core enzyme.



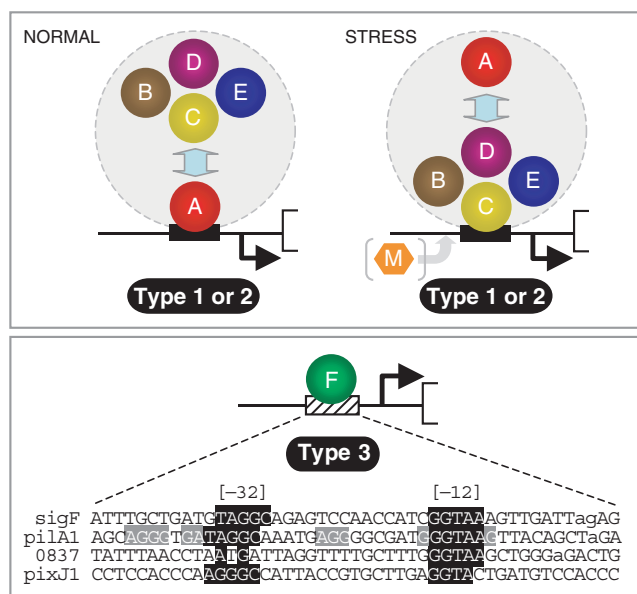
**Figure 6.** Identification of novel promoters recognized by SigF. Primer extension was performed with the SLL0837-R (5'-ACGT TTTGACACAGTAAATTTTTC-3') or SLL0041-R (5'-GTCAGGACGCGCATCGGAGGAAGG-3') primer and total RNA (10  $\mu$ g each) prepared from 6803 WT,  $\Delta F$  or  $\Delta F$  complemented with *sigF* (C1, Figure 2), respectively. Signals correspond to transcription start points (+1) of Psl0837-41 or PpixJ1-32. A molecular size marker as a sequencing ladder was synthesized with the plasmid DNA of pPILA76\_puc and the RV primer. The X-ray film was subjected to autoradiography for 3 days.

### Identification of novel promoters recognized by SigF

The core promoter sequence of PpilA1-54 was developed. Similarity to the core motif was also observed upstream of the PCC 6803 *sigF* gene. To identify novel gene promoters recognized by SigF in PCC 6803, we searched for the conserved sequence GGGTAAAG, GGGT[A/G], [C/G]GT[A/G][A/G/T] or [C/T]AGGC [N10-30] GGGT[A/G][A/G][A/G] within 200 bp upstream of genes in the PCC 6803 genome. We found more than 50 genes and chose several for a primer extension analysis as in Figure 2C. The results using specific primers are shown in Figure 6. The transcripts of *sll0837* (periplasmic protein of unknown function) and *sll0041* (phytochrome-like phototaxis protein, histidine kinase of a two-component system, *pixJ1* = *pisJ1* = *taxD1*, *tsr* or *cheD* homologue) were present at very low levels or had disappeared in the  $\Delta F$  strain. They were also compensated in the C1 strain, clearly indicating that their transcription depended on SigF.

### DISCUSSION

This is the first report characterizing the promoter sequence recognized by a group 3  $\sigma$ -factor in photosynthesizing eubacteria. The region containing the core promoter sequence of PCC 6803 *pilA1* is presented in Figure 7 (bottom). Sequences in the  $-32$  (TAGGC) and  $-12$  (GGTAA) regions of *pilA1* are similar to that of PCC 6803 *sigF*. Similar promoter sequences were also observed upstream of other *pilA* genes in cyanobacteria, *Gloeobacter violaceus* PCC 7421, *Thermosynechococcus* BP-1, *Anabaena* PCC 7120 and *Synechococcus* PCC 6301 (data not shown). In addition, the promoter region recognized by SigF may be AG-rich. For example, the AG content is 71% (32 bp/45 bp) in the region from  $-44$  to  $+1$  ( $+1$  as a transcription start point) of PCC 6803 *pilA1*, the sequences of which involve purine tracts such as AAA ( $-29$  to  $-27$ ), AAG ( $-11$  to  $-9$ ), AGG ( $-41$  to  $-39$ ,  $-33$  to  $-31$ ,  $-24$  to  $-22$ ) and GGG ( $-23$  to  $-20$ ,  $-15$  to  $-13$ ). These tracts in (and out of) the core promoter sequence seem to be important for transcription *in vitro* (Figure 4). Thus, a long AG-rich sequence possessing the consensus



**Figure 7.** Promoter types of PCC 6803. Type 1 and 2 promoters (black boxes for  $-35$  and/or  $-10$  sequences) and RNAP  $\sigma$ -factors (color circle) are shown. A transcriptional modulator protein (M) is presented as a hexagon. Coexistence and redundant promoter recognition by group 1 and 2  $\sigma$ -factors are indicated in gray circles. Effective AG sequences (based on the results of Figure 4) or the core sequences conserved in type 3 promoters are indicated in gray or black. Other details are described in the text.

motif at  $-32$  and  $-12$  is characteristic of the PCC 6803 *pilA1* promoter that is particularly distinct from the consensus promoter sequence in the  $-35$  (TTGACA) and  $-10$  (TATAAT) regions in cyanobacteria (6,21,22,29). Of note, the PpilA1-54 region has a set of sequences at  $-31$  (TAGGCA) and  $-9$  (TAAGTT) that exhibits a low degree of similarity to that of the cyanobacterial consensus promoter, mentioned above. However, this sequence does not function *in vivo* and *in vitro* as the *E. coli*  $\sigma^{70}$ -type promoter (Figures 2, 3 and 5). Moreover, the sequence in  $-12$  the region may be more significant since PpilA1-54 transcription was relatively very limited in the mutants of mt3-No9 to No11 and mt7-NoVI (Figure 4). In fact, a high level of similarity of the GGTA sequence is conserved at the  $-12$  region in the promoters recognized by SigF (Figure 7).

Furthermore, the proof reading of RNAP transcription seems to need not only SigF but also the function of the cyanobacterial core enzyme (Figure 5C). The  $\beta'$  subunit is usually split into two parts in cyanobacteria,  $\gamma$  (RpoC1) and  $\beta'$  (RpoC2), the same as in higher plant chloroplasts. Thus, the RNAP holoenzyme of cyanobacteria consists of a core enzyme with the subunit structure  $\alpha_2\beta\beta'\gamma$  and one of the species of  $\sigma$ -factor (31). The eubacterial  $\beta'$  (RpoC) subunit, for example of *E. coli*, possesses eight highly conserved domains, A to H (32). In cyanobacteria and also plant chloroplasts, RpoC1 and RpoC2 possess the domains A to D and E to H, respectively (33,34). In addition, RpoC2 has  $\sim 50$ – $70$  kDa insert roughly in the middle between domains G and H (10). This insert may affect the

global architecture and potential interactions of the holoenzyme and can easily account at least for some of the differences observed between *E. coli* and cyanobacteria in behavior towards given promoter sequences (10,35).

On the other hand, the AG content is 56% (25 bp/45 bp) in the region from  $-44$  to  $+1$  of PCC 6803 *sigF*, lower than that of *pilA1* (Figure 7). The low levels of AG nucleotides in the *sigF* promoter region might correlate to the weaker expression of the *sigF* transcript than *pilA1* transcript in PCC 6803 cells (Figure 2C). The AG content is 47% (21 bp/45 bp) in the region from  $-44$  to  $+1$  of *sll0837*, lower than that of *sigF*. The  $-32$  region of Psl0837-41 exhibits a low degree of similarity to the TAGGC sequence (Figure 7). The AG content is 42% (19 bp/45 bp) in the region from  $-44$  to  $+1$  of *sll0041* (*pixJ1* = *taxD1*), lower than that of *sll0837*. The GGTA sequence of PpixJ1-32 was mapped in the  $-20$  region and is slightly off from the  $-12$  region, suggesting a unique architecture. Actually, weak expression was observed for the transcripts from Psl0837-41 and PpixJ1-32, and a longer exposure time was required for the detection of signals (Figure 6) than in the case of PpilA1-54 and PsigF-22/-21 (Figure 2). The low sequence similarity in the core promoter of Psl0837-41 and PpixJ1-32 encompassing unique structures may cause such weak expression in the cells. In addition, not only the unique sequence but also the promoter architecture may bring out the stringent promoter recognition by SigF, since the transcription depends on template DNA forms (Figure 3B).

The *pilA1* promoter is recognized by only SigF  $\sigma$ -factor in PCC 6803 (Figure 2A and B). The *sigF* knockout strain is still viable even with the loss of pilus formation and motility (6,7,18). This suggests that SigF does not recognize promoters of indispensable genes for cell viability but for fine tuning cell maintenance. SigF was assigned as a group 3  $\sigma$ -factor but might not be the ECF type in PCC 6803 by analogy of amino acid sequences (6,9,18,30). In *E. coli* or *Pseudomonas aeruginosa* (Gram-negative bacterium), RpoF ( $\sigma^{28}$ , Figure 1A) can recognize the promoter, TAAA-N15-GCCGAWAA-N4-start ( $+1$ ), related to chemotaxis, flagellin and motility genes for cell surface structure (9,36). The type of promoter sequence seems to have little or no similarity to those recognized by PCC 6803 SigF. In fact, the PCC 6803 *pilA1* promoter was not recognized by *E. coli*  $\sigma$ -factors in the cell (Figure 2D). In *B. subtilis* (Gram-positive bacterium), SigD ( $\sigma^{28}$ ) is assigned as the *E. coli* RpoF type (Figure 1A) and can also recognize the promoter, TAAA-N15-GCCGATAT, of chemotaxis, autolysin and flagellin genes (37,38). However, 6803 SigF is rather assigned into another node in the phylogenetic tree (Figure 1A) involving a  $\sigma$ -factor, SigB ( $\sigma^{37}$ , Figure 1A) that contributes to the general stress response in *B. subtilis*. This notion also accords to our previous sequence analogy in which the estimated PCC 6803 *sigF* promoter partially exhibited similarity to the sequence recognized by *B. subtilis* SigB (6). Bs\_SigB ( $\sigma^{37}$ ) can recognize a promoter, RGGXTTRA-N14-GGGTAT, the sequence of which is partially similar to that at the  $-12$  promoter of PCC 6803 SigF (Figure 7). These findings strongly suggest that the structure and/or

promoter selectivity of PCC 6803 SigF is novel for eubacterial group 3  $\sigma$ -factors.

In this study, two novel genes, the promoters of which might be recognized by SigF, were identified. Sll0837 is a homologue of a periplasmic protein of unknown function. Sll0041 is a phytochrome-like protein (PixJ1 = PisJ1 = TaxD1), having a domain of signal transduction histidine kinase as a two-component system, for positive phototaxis and is also homologous to methyl-accepting chemotaxis proteins (CheD and Tsr) or the type IV pilus biogenesis protein PilJ (39–41). These findings demonstrate that SigF can recognize the promoters of genes for cell surface-dependent phenomena.

PCC 6803 has four kinds of group 3  $\sigma$ -factors, SigF, SigG, SigH and SigI, whose molecular masses (23–30 kDa) deduced from the amino acid sequences are lower than those of SigA (61 kDa) as a group 1 factor, and SigB, SigC, SigD and SigE (38–52 kDa) as group 2  $\sigma$ -factors. The functions of SigG, SigH and SigI have not been clarified in PCC 6803. Studies of the functional domain in PCC 6803 SigF will be required to elucidate the mechanism of stringent promoter selectivity.

PCC 6803 promoter types connected with RNAP are summarized below, based on the results of this study and previous reports (Figure 7). Type 1 promoters possess both –35 (TTGACA) and –10 (TATAAT) hexamers as the consensus sequence. For example, the light-responsive expression of *psbA* has been characterized in PCC 6803 (6,9,10,17,24). The group 1  $\sigma$ -factor SigA can constitutively recognize the *psbA* promoter for basal transcription (normal condition), and the group 2  $\sigma$ -factor SigD contributes to the light-induced transcription of *psbA* under light (9,17). In addition, the group 2  $\sigma$ -factor SigE (see Introduction section) also contributes cooperatively to light-induced transcription on the *psbA* promoter (17). Specific transcription has been observed *in vitro* by RNAP with group 1 and 2  $\sigma$ -factors (10), indicating that they coexist and redundantly recognize the type 1 promoter, changing their species, population and/or activity under stress conditions (6,10,13,17,18,30). The type 2 promoter possesses only the –10 hexamer, and an enhancer sequence associated with a transcriptional activator protein compensates for the lack of the –35 hexamer. For instance, the PCC 6803 *glnB* P2 promoter (PglNB-54/-53) has only the –10 hexamer, and NtcA (hexagonal form in Figure 7) binds to a consensus sequence motif of GTA-N8-TAC, upstream of the promoter, resulting in transcriptional enhancement (12,13). In the report, SigB and SigC (or SigE) contribute to the transcription from PglNB-54/-53 with a  $\sigma$ -factor replaced in a growth-phase-dependent manner under nitrogen-deprived (stress) conditions. They also indicate that multiple group 2  $\sigma$ -factors take part in the transcription from the type 2 promoter in cooperation with the group 1  $\sigma$  factor, SigA (12,13,30). The type 3 promoter is distinct from type 1 and type 2 promoters and the transcription may be independent of group 1 and group 2  $\sigma$  factors, as with the *pilA1* promoters.

It remains to be studied whether all group 3  $\sigma$ -factors contribute to transcription on type 3 promoters. At least, previous data suggested that 6803 SigF can not recognize

*in vitro* 6803 *psbA* promoter, assigned as the type 1 promoter (6,10). In this study, SigF could recognize *in vitro* mutagenized PpilA1-54 (mt22) of whose sequence was changed to the type 1 promoter (Figure 5B). However, it might be that mutagenized PpilA1-54 (mt22) still has a partially conserved sequence recognized by SigF. The influence of expression at the transcript and protein levels among group 1, 2 and 3  $\sigma$ -factor genes in PCC 6803 has been reported recently (18,42,43). Further identification and characterization of other promoters recognized by SigF will help to clarify this issue. Analyses using an *in vitro* transcription system together with knockout strains may also be useful for the identification and characterization of novel promoters recognized by other group 3  $\sigma$ -factors, SigG, SigH and SigI.

## ACKNOWLEDGEMENTS

The authors thank Dr V.V. Zinchenko and Dr I. Suzuki for the gifts of pVZ and R751 plasmids. They also thank Dr A. Taton for the computer analysis of the promoter and Dr K. Tanaka's group for providing  $\Delta D$  and  $\Delta E$  strains. The authors also thank Dr T. Nishizawa for technical advice and T. Yoshimura for pVZ321SIGF-COMP. This work was supported in part by grants from Ibaraki University, The Foundation for Earth Environment, The New Technology Development Foundation and Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan to M.A. Funding to pay the Open Access publication charges for this article was provided by M.A.

*Conflict of interest statement.* None declared.

## REFERENCES

- Ishihama, A. (1993) Protein-protein communication within the transcription apparatus. *J. Bacteriol.*, **175**, 2483–2489.
- Ishihama, A. (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annu. Rev. Microbiol.*, **54**, 499–518.
- Wösten, M.M. (1998) Eubacterial sigma-factors. *FEMS Microbiol. Rev.*, **22**, 127–150.
- Lonetto, M., Gribskov, M. and Gross, C.A. (1992) The  $\sigma^{70}$  family: sequence conservation and evolutionary relationships. *J. Bacteriol.*, **174**, 3843–3849.
- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S. *et al.* (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, **3**, 109–136.
- Imamura, S., Yoshihara, S., Nakano, S., Shiozaki, N., Yamada, A., Tanaka, K., Takahashi, H., Asayama, M. and Shirai, M. (2003) Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium. *J. Mol. Biol.*, **325**, 857–872.
- Bhaya, D., Watanabe, N., Ogawa, T. and Grossman, A.R. (1999) The role of an alternative sigma factor in motility and pilus formation in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Proc. Natl Acad. Sci. USA*, **96**, 3188–3193.
- Huckauf, J., Nomura, C., Forchhammer, K. and Hagemann, M. (2000) Stress responses of *Synechocystis* sp. strain PCC 6803 mutants impaired in genes encoding putative alternative sigma factors. *Microbiology*, **146**, 2877–2889.
- Imamura, S., Asayama, M., Takahashi, H., Tanaka, K. and Shirai, M. (2003) Antagonistic dark/light-induced SigB/SigD, group 2 sigma



- factors, expression through redox potential and their roles in cyanobacteria. *FEBS Lett.*, **554**, 357–362.
10. Imamura, S., Asayama, M. and Shirai, M. (2004) *In vitro* transcription analysis by reconstituted cyanobacterial RNA polymerase: roles of group 1 and 2 sigma factors and a core subunit, RpoC2. *Genes Cells*, **9**, 1175–1187.
  11. Tuominen, I., Pollari, M., Tyystjärvi, E. and Tyystjärvi, T. (2006) The SigB sigma factor mediates high-temperature responses in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett.*, **580**, 319–323.
  12. Asayama, M., Imamura, S., Yoshihara, S., Miyazaki, A., Yoshida, N., Sazuka, T., Kaneko, T., Ohara, O., Tabata, S., Osanai, T. *et al.* (2004) SigC, the group 2 sigma factor of RNA polymerase, contributes to the late-stage gene expression and nitrogen promoter recognition in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Biosci. Biotechnol. Biochem.*, **68**, 477–487.
  13. Imamura, S., Tanaka, K., Shirai, M. and Asayama, M. (2006) Growth phase-dependent activation of nitrogen-related genes by a control network of group 1 and group 2  $\sigma$  factors in a cyanobacterium. *J. Biol. Chem.*, **281**, 2668–2675.
  14. Muro-Pastor, A. M., Herrero, A. and Flores, E. (2001) Nitrogen-regulated group 2 sigma factor from *Synechocystis* sp. strain PCC 6803 involved in survival under nitrogen stress. *J. Bacteriol.*, **183**, 1090–1095.
  15. Osanai, T., Kanesaki, Y., Nakano, T., Takahashi, H., Asayama, M., Shirai, M., Kanehisa, M., Suzuki, I., Murata, N. and Tanaka, K. (2005) Positive regulation of sugar catabolic pathways in the cyanobacterium *Synechocystis* sp. PCC 6803 by the group 2  $\sigma$  factor sigE. *J. Biol. Chem.*, **280**, 30653–30659.
  16. Osanai, T., Imamura, S., Asayama, M., Shirai, M., Suzuki, I., Murata, N. and Tanaka, K. (2006) Nitrogen induction of sugar catabolic gene expression in *Synechocystis* sp. PCC 6803. *DNA Res.*, **13**, 185–195.
  17. Yoshimura, T., Imamura, S., Tanaka, K., Shirai, M. and Asayama, M. (2007) Cooperation of group 2  $\sigma$  factors, SigD and SigE for light-induced transcription in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett.*, **581**, 1495–1500.
  18. Matsui, M., Yoshimura, T., Wakabayashi, Y., Imamura, S., Tanaka, K., Takahashi, H., Asayama, M. and Shirai, M. (2007) Interference expression at levels of the transcript and protein among group 1, 2, and 3 sigma factor genes in a cyanobacterium. *Microbes Environ.*, **22**, 32–43.
  19. Agrawal, G.K., Kato, H., Asayama, M. and Shirai, M. (2001) An AU-box motif upstream of the SD sequence of light-dependent *psbA* transcripts confers mRNA instability in darkness in cyanobacteria. *Nucleic Acids Res.*, **29**, 1835–1843.
  20. Asayama, M., Hayasaka, Y., Kabasawa, M., Shirai, M. and Ohyama, A. (1999) An intrinsic DNA curvature found in the cyanobacterium *Microcystis aeruginosa* K-81 affects the promoter activity of *rpoD1* encoding a principal sigma factor. *J. Biochem.*, **125**, 460–468.
  21. Shibato, J., Asayama, M. and Shirai, M. (1998) Specific recognition of the cyanobacterial *psbA* promoter by RNA polymerases containing principal sigma factors. *Biochim. Biophys. Acta*, **1442**, 296–303.
  22. Shibato, J., Agrawal, G.K., Kato, H., Asayama, M. and Shirai, M. (2002) The 5'-upstream cis-acting sequences of a cyanobacterial *psbA* gene: analysis of their roles in basal, light-dependent and circadian transcription. *Mol. Genet. Genomics*, **267**, 684–694.
  23. Li, R. and Golden, S.S. (1993) Enhancer activity of light-responsive regulatory elements in the untranslated leader regions of cyanobacterial *psbA* genes. *Proc. Natl Acad. Sci. USA*, **90**, 11678–11682.
  24. Horie, Y., Ito, Y., Ono, M., Moriwaki, N., Kato, H., Hamakubo, Y., Amano, T., Wachi, M., Shirai, M. and Asayama, M. (2007) Dark-induced mRNA instability involves RNase E/G-type endoribonuclease cleavage at the AU-box and SD sequences in cyanobacteria. *Mol. Genet. Genomics*, **278**, 331–346.
  25. Meyer, R.J. and Shapiro, J.A. (1980) Genetic organization of the broad-host-range IncP-1 plasmid R751. *J. Bacteriol.*, **143**, 1362–1373.
  26. Zinchenko, V.V., Piven, I.V., Melnik, V.A. and Shestakov, S.V. (1999) Vectors for the complementation analysis of the cyanobacterial mutants. *Russian J. Genet.*, **35**, 228–232.
  27. Asayama, M., Tanaka, K., Takahashi, H., Sato, A., Aida, T. and Shirai, M. (1996) Cloning, sequencing and characterization of the gene encoding a principal sigma factor homolog from the cyanobacterium *Microcystis aeruginosa* K-81. *Gene*, **181**, 213–217.
  28. Asayama, M., Kato, H., Shibato, J., Shirai, M. and Ohyama, T. (2002) The curved DNA structure in the 5'-upstream region of the light-responsive genes: its universality, binding factor and function for cyanobacterial *psbA* transcription. *Nucleic Acids Res.*, **30**, 4658–4666.
  29. Ito, Y., Asayama, M. and Shirai, M. (2003) Light-responsive *psbA* transcription requires the -35 hexamer in the promoter and its proximal upstream element, UPE, in cyanobacteria. *Biosci. Biotechnol. Biochem.*, **67**, 1382–1390.
  30. Asayama, M. (2006) Regulatory system for light-responsive gene expression in photosynthesizing bacteria: cis-elements and trans-acting factors in transcription and post-transcription. *Biosci. Biotechnol. Biochem.*, **70**, 565–573.
  31. Schneider, G.J., Tumer, N.E., Richaud, C., Borbely, G. and Haselkorn, R. (1987) Purification and characterization of RNA polymerase from the cyanobacterium *Anabaena* 7120. *J. Biol. Chem.*, **262**, 14633–14639.
  32. Allison, L.A., Moyle, M., Shales, M. and Ingles, C.J. (1985) Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell*, **42**, 599–610.
  33. Schneider, G.J. and Haselkorn, R. (1988) RNA polymerase subunit homology among cyanobacteria, other eubacteria and archaeobacteria. *J. Bacteriol.*, **170**, 4136–4140.
  34. Bergsland, K.J. and Haselkorn, R. (1991) Evolutionary relationships among eubacteria, cyanobacteria, and chloroplasts: evidence from the *rpoC1* gene of *Anabaena* sp. strain PCC 7120. *J. Bacteriol.*, **173**, 3446–3455.
  35. Schyns, G., Jia, L., Coursin, T., Tandeau de Marsac, N. and Houmard, J. (1998) Promoter recognition by a cyanobacterial RNA polymerase: *in vitro* studies with the *Calothrix* sp. PCC 7601 transcriptional factors RcaA and RcaD. *Plant Mol. Biol.*, **36**, 649–659.
  36. Totten, P.A. and Lory, S. (1990) Characterization of the type a flagellin gene from *Pseudomonas aeruginosa* PAK. *J. Bacteriol.*, **172**, 7188–7199.
  37. Helmann, J.D. (1991) Alternative sigma factors and the regulation of flagellar gene expression. *Mol. Microbiol.*, **5**, 2875–2882.
  38. Haldenwang, W.G. (1995) The sigma factors of *Bacillus subtilis*. *Microbiol. Rev.*, **59**, 1–30.
  39. Yoshihara, S., Suzuki, F., Fujita, H., Geng, X.X. and Ikeuchi, M. (2000) Novel putative photoreceptor and regulatory genes required for the positive phototactic movement of the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.*, **41**, 1299–1304.
  40. Bhaya, D., Takahashi, A. and Grossman, A.R. (2001) Light regulation of type IV pilus-dependent motility by chemosensor-like elements in *Synechocystis* PCC6803. *Proc. Natl Acad. Sci. USA*, **98**, 7540–7545.
  41. Yoshihara, S. and Ikeuchi, M. (2004) Phototactic motility in the unicellular cyanobacterium *Synechocystis* sp. PCC6803. *Photochem. Photobiol. Sci.*, **3**, 512–518.
  42. Lemeille, S., Latifi, A. and Geiselmann, J. (2005) Inferring the connectivity of a regulatory network from mRNA quantification in *Synechocystis* PCC6803. *Nucleic Acids Res.*, **33**, 3381–3389.
  43. Summerfield, T.C. and Sherman, L.A. (2007) Role of sigma factors in controlling global gene expression in light/dark transitions in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.*, **189**, 7829–7840.