Functional Characterization of a Cyanobacterial OmpR/PhoB Class Transcription Factor Binding Site Controlling Light Color Responses

Ryan P. Bezy and David M. Kehoe*

Department of Biology, 1001 East Third Street, Indiana University, Bloomington, Indiana 47405

Received 26 May 2010/Accepted 26 July 2010

Cyanobacteria possess the ability to modify their photosynthetic light-harvesting antennae, or phycobilisomes (PBS), in response to changes in a wide range of abiotic cues, including both light intensity and color (20–22). Color responsiveness is called chromatic acclimation. This process is widespread in marine and freshwater environments and exists in multiple forms, the best characterized of which is called complementary chromatic acclimation (CCA; historically referred to as “adaptation”) (15, 30, 38, 44). The model organism for the study of CCA and its regulation is the freshwater filamentous cyanobacterium *Fremyella diplosiphon* UTEX 481/Tolyphothrix sp. PCC 7601. During CCA, PBS production in this organism is influenced by the ratio of green light (GL) to red light (RL) in the environment. Growth in RL leads to the accumulation of two different forms of an RL-absorbing chromoprotein called phycocyanin (PC) in the outer regions of the PBS. These two forms are called PC1 and PC2. Growth in GL results in the accumulation of PC1 and another chromoprotein, called phycocyrtin (PE), which maximally absorbs GL, in the outer portions of the PBS. Thus, cyanobacteria that are capable of CCA can maximize photon capture for photosynthesis by tailoring the color absorption profiles of their PBS to the spectral distribution of ambient light in the green and red regions of the spectrum.

CCA-controlled changes in PC2 and PE abundance have so far been found to occur entirely at the level of transcription. Transcripts from the *cpcB2A2H2I2D2* operon (hereafter called the *cpc2* operon), which encodes the PC2 apoproteins as well as PBS structural proteins called linkers, are highly abundant in RL and virtually undetectable in GL (10, 12, 34). Conversely, RNA levels from the *cpcB4* operon, which encodes the apo-PE subunits, and the *cpeCDESTR* operon (hereafter called the *cpeC* operon), whose first three genes encode the PE linkers, are high in GL and low in RL (16–18, 36). Many additional genes are CCA regulated, including *pcyA*, which encodes the biosynthetic enzyme necessary to produce the chromophore used for RL absorption by PC and is expressed five times more highly in RL than in GL in *F. diplosiphon* (1). There are also many genes that are either weakly or not CCA regulated, including a number that encode components of the PBS that are present in both RL and GL, such as the operon encoding PC1, *cpcB1A1* (herein called the *cpc1* operon) (11, 37, 43).

CCA is regulated by two distinct signal transduction pathways (30). The best characterized of these is the Rca system, which is a complex two-component system. It consists of a sensor called RcaE that contains a chromophore binding domain and is a member of the phytochrome superfamily (28, 45), a single-domain response regulator named RcaF that appears to act after RcaE (27), and a multidomain response regulator called RcaC that is a member of the OmpR/PhoB family of transcription factors and apparently acts after RcaF (7, 27). The Rca system controls both RL- and GL-upregulated genes by the binding of RcaC to a 7-bp direct repeat DNA sequence, separated by 4 bp, called the L box (5′-TTGCACA N₇TTGCACA-3′) (32). Each of the promoters of the genes that are upregulated in RL, i.e., *pcyA* and the *cpc2* operon, contains an L box, with the 3′ end located at positions −22 and −24, respectively, relative to the CCA-regulated transcription start site, and these function as activating elements in RL (1) (Fig. 1A). There is also an L box upstream of the *cpeC* operon,
but it is in the opposite orientation from that of the pcyA and cpc2 L boxes, and its 3′ end is located at position −71 (1). Thus far, no other GL-upregulated operons have been found to be flanked by an L box, consistent with the fact that their regulation by the Rca system appears to be through the cpeC operon, since the final gene in that operon encodes an activator called CpeR that is required for the expression of cpeBA operon and other GL-upregulated genes (2, 8, 40). The cpeC L box was recently shown to function as a negative regulator of transcription in RL (32). Thus, in RL the L box acts as both an activating and a repressing element, and these activities have been correlated with the orientation and location of this element relative to the transcription start sites of the genes it controls.

In addition to the ability of L boxes associated with different genes to either activate or repress transcription, each specific L box imparts a different relative degree of induction or repression of light-regulated gene expression. The cpe2 L box mediates a ≥20-fold increase in expression in RL compared to the GL level (6), the pcyA L box imparts 5-fold induction in RL (1), and although the differences in location and orientation make comparisons to the other two L boxes difficult, the cpeC L box provides 2- to 3-fold repression of expression in RL (32). Interestingly, the apparent binding affinity of RcaC for each of these L boxes mirrors these differences in activity, since it is strongest for the cpe2 operon and weakest for the cpeC operon (32). These differences in activity and binding may be due to variations in the sequences within the two repeats or flanking them: the repeats are identical in the cpe2 and pcyA L boxes but contain a 1-bp change in the cpeC L box, while the sequences between and flanking the pcyA L box repeats are quite diverged from those found around both the cpe2 and cpeC repeats, which are very highly conserved (1).

In addition to the issue of how local sequence variations affect L box activity levels, there is also the question of how these elements function as both activating and repressing elements under the same physiological conditions. The most likely possibility is that the differences are controlled by the position and orientation of these elements relative to the transcription start site. It is also very possible that the activities of some or all of these L boxes are modified by other DNA sequence elements and their binding proteins, although at least for the cpe2 L box, the sequences within the region approximately 250 bp upstream of the L box do not further boost activity in RL (33), and although the reason(s) for these differences in promoter strength is not yet understood, it is clear that the strength of these promoters in the absence of Rca system influence predisposes them to control via repression or activation.

The study of RcaC and L box function is important because OmpR/PhoB class transcription factors are abundant in cyanobacteria (32), but unlike the case for other bacteria (31), we have little knowledge about the mechanisms through which they operate. Here we provide a functional analysis of the effects of changing the L box sequence, position, and orientation on the ability of this element to both activate and repress transcription via the Rca regulatory pathway.

FIG. 1. Locations and sequences of L boxes in the genomes of *Fremyella diplosiphon* and two other cyanobacteria. (A) Placement and orientation of known L boxes (closed triangles) within the *F. diplosiphon*, *Synechococcus* sp. PCC 7335, and *Pseudanabaena* sp. PCC 7409 genomes. The orientation of each triangle denotes the directionality of the L box in front of each open reading frame, indicated with arrows. (B) Alignment of L box sequences upstream of pcyA and the cpe2 and pebAB operons in *Synechococcus* sp. PCC 7335 (Syn), pcyA and the cpe2 and cpeC operons in *F. diplosiphon* (Fd), and the cpe2 operon in *Pseudanabaena* sp. PCC 7409 (Psa). L box repeats are underlined. Shaded letters indicate shared sequence identity with the consensus L box repeat (see the text). Nonitalicized numbers are for known transcription start sites in *F. diplosiphon* and *Pseudanabaena* sp. PCC 7409, and italicized numbers are for presumptive translation start sites in *Synechococcus* sp. PCC 7335. The dotted line marks the sequence that makes up the L box repeats and flanking sequences within the *F. diplosiphon* (F. dip.) between *F. diplosiphon* and *Synechococcus* sp. PCC 7335 (7335), and within *Synechococcus* sp. PCC 7335.
TABLE 1. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Construct</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCc3’B</td>
<td>GCGGGATCCTGCTTGAGGAGCTACTTTTGGCC</td>
<td>(BamHI site underlined)</td>
</tr>
<tr>
<td>PCc5’P</td>
<td>GCGGTGTAGAGATTTACAAATCTTGATGTACA (SphI site underlined)</td>
<td></td>
</tr>
<tr>
<td>PC1Gjxn</td>
<td>GCGGGATCCTGCTTGAGGAGCTACTTTTGGCC (BamHI site underlined)</td>
<td></td>
</tr>
<tr>
<td>400cpeC</td>
<td>GCGGATATGGCATCTGGCGAATCTTGAGATGGCTGTG    (SphI site underlined)</td>
<td></td>
</tr>
</tbody>
</table>

Primers for cpeC promoter truncations

200cpeC pRB2 GCGGATGTTCTGGGATGAGGAGGATAATGGG (SphI site underlined)  
150cpeC pRB3 GCGGTGTAGAGATTTACAAATCTTGATGTACA (SphI site underlined)  
110cpeC pRB4 GCGGATGTTCTGGGATGAGGAGGATAATGGG (SphI site underlined)  
52cpeC pRB5 GCGGATATGGCATCTGGCGAATCTTGAGATGGCTGTG    (SphI site underlined)  

Primers for cpeC substitutions (−150 to −90 region)

130DRF pRB1DA ACTAGTCTATCCGGATGCTACTAAACAGCTGCTATGAATGG  
130DRR pRB1DB GCGGTGTAGAGATTTACAAATCTTGATGTACA (SphI site underlined)  
InvRepF pRB1B ATTACCGAGATCCACTGGAAGCTGATACATAGTTTGTGCAATTAGGTGCAAA  
InvRepR pRB1DB GCGGTGTAGAGATTTACAAATCTTGATGTACA (SphI site underlined)  
NearlF pRB1DA CTCCTGCGTTATCTGCTGCAATCTGGAATGTTGTGCAAAATTCCTCGTCC  
NearlR pRB1DB CCAATGCTGAGTAGTGCGAGATGGGAAATGTTGGAATTGGCAAGTCG  
130inv-Dbl pRB1AB GGGAATACATTACATCCTTCCCAATTACCCATTACCAGATCCAGTGACATA  
130DRR pRB1DB GCGGATATGGCATCTGGCGAATCTTGAGATGGCTGTG    (SphI site underlined)  

Primers for cpc1 L box variation

PCc-LBrand1 pPCc1 TCAGGGTATGATTTTTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-LBrand2 pPCc1 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-cpc2-LB1 pPCc2 GTTGAGGTGAGCGGAGATGCTTATTTTCTGCTATGAAAGCTTG  
PCc-cpc2-LB2 pPCc2 GCTTTGGTCACTAATCAATCAATAGGCTTATTTTCTGCTATGAAAGCTTG  
PCc-33cpeC2 pPCc10 GAATTTTGCACTCAATTTGCACACAACTACGATATAGTATAAACAAGTAATGGGGAAG  
PCc-33cpeC1 pPCc10 TCAATTTATGAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-cpc2-LB2 pPCc2 GCTTTGGTCACTAATCAATCAATAGGCTTATTTTCTGCTATGAAAGCTTG  
PCc-cpc2-LB1 pPCc2 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-cpc2-5lb2 pPCc3 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-cpc2-5lb1 pPCc3 TGAATTTTGGAATTTTTGCAAAAATACTACCAATACGATATAGTATAAACAAGTAATGGGGAAG  

PCc-LBrand2 pPCc1 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-LBrand1 pPCc1 TCAGGGTATGATTTTTGACATTTAGTCTTGCTATGAAAGCTTG  

PCc-pcya5lb2 pPCc5 GCGGATGTTGACTATCCGGATGCTACTAAACAGCTGCTATGAATGG  
PCc-pcya5lb1 pPCc5 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-pcyalb2 pPCc4 GCGGATGTTGACTATCCGGATGCTACTAAACAGCTGCTATGAATGG  
PCc-pcyalb1 pPCc4 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-cpc2-5lb2 pPCc3 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  
PCc-cpc2-5lb1 pPCc3 TGAATTTTGGAATTTTTGCAAAAATACTACCAATACGATATAGTATAAACAAGTAATGGGGAAG  
PCc-cpc2lb2 pPCc2 GCTTTGGTCACTAATCAATCAATAGGCTTATTTTCTGCTATGAAAGCTTG  
PCc-cpc2lb1 pPCc2 TAAATTTAGTGACATTTAGTCTTGCTATGAAAGCTTG  

Plasmid construction. All position numbers used are relative to the transcriptional start site unless otherwise noted. All PCR-amplified DNA regions and ligation junctions were checked by sequencing to ensure that no mutations had occurred during cloning.

To create a series of 5’ cpc1 promoter truncations (Table 1), the cpc1 promoter that created different versions of the L box had to be reassembled. Mismatched sequences into the cpeC promoter that created different versions of the L box and provided a region of sequence overlap between the upstream and downstream amplified fragments that allowed annealing and a final PCR amplification using the outside primers from the first two amplifications. This created a series of cpeC promoters from positions −412 to +286 with altered L box sequences, as described below. The upstream and downstream outside primers also added PstI and BamHI sites, respectively, which were used to cut the final PCR amplification product for insertion into similarly cut pCC412G (6). Primer sequences and the plasmids they were used to create are listed in Table 1.

pPC2-C was made by PCR amplifying the cpc2 promoter from positions −285 to +20, using the PCiPR and PCiPL primers. The product was digested with SphI and PstI and inserted into similarly cut p400pcC (32). The name of plasmid p400pcC was changed to PRBl for simplicity.

To create a series of 5’ cpc1 promoter truncations (Table 1), the cpc1 promoter and cpeC 5’ leader were PCR amplified, using pRB1 as a template, a series of upstream primers corresponding to sequences at truncation endpoints, and the downstream primer PC1Gjxn. The upstream primers contained an SphI site, mutated sequences into the cpeC promoter that created different versions of the L box and provided a region of sequence overlap between the upstream and downstream amplified fragments that allowed annealing and a final PCR amplification using the outside primers from the first two amplifications. This created a series of cpeC promoters from positions −412 to +286 with altered L box sequences, as described below. The upstream and downstream outside primers also added PstI and BamHI sites, respectively, which were used to cut the final PCR amplification product for insertion into similarly cut pCC412G (6). Primer sequences and the plasmids they were used to create are listed in Table 1.

pPC2-C was made by PCR amplifying the cpc2 promoter from positions −285 to +20, using the PCiPR and PCiPL primers. The product was digested with SphI and PstI and inserted into similarly cut p400pcC (32). The name of plasmid p400pcC was changed to PRBl for simplicity.

To create a series of 5’ cpc1 promoter truncations (Table 1), the cpc1 promoter and cpeC 5’ leader were PCR amplified, using pRB1 as a template, a series of upstream primers corresponding to sequences at truncation endpoints, and the downstream primer PC1Gjxn. The upstream primers contained an SphI site, 

**MATERIALS AND METHODS**

**Strains, growth conditions, and transformations.** The shortened filament mutant strain SF33 (9) of *Fremyella diplosiphon* UTEX 481 (also known as *Tolyphora fuscata* sp. PCC 7801) was the wild type, and cultures were grown as previously described (40), with or without 10 μg of kanamycin (Kan) per ml. Cells were grown in light intensities of approximately 15 μmol photons m⁻² s⁻¹ and transformed by electroporation as described previously (29).

**Plasmid construction.** All position numbers used are relative to the transcriptional start site unless otherwise noted. All PCR-amplified DNA regions and ligation junctions were checked by sequencing to ensure that no mutations had occurred during cloning.

Constrasts with the cpc1 promoter harboring various versions of the L box were created by separate PCR amplifications of the upstream and downstream regions. An outside primer, PCc5’P, was paired with one of several inside primers to amplify the upstream portion. A second outside primer, PCc3’B, was paired with one of several other inside primers to amplify the downstream portion of the promoter and the 5’ leader of cpeC. The inside primers introduced

---

*Vol. 192, 2010 CYANOBACTERIAL L BOX CHARACTERIZATION*
and PC1Gjn contained a BamHI site. Amplification products were cut with Splh and BamHI and inserted into similarly cut pRB1. Mutations in the cpeC promoter between positions −150 and −90 (Table 1) were created via two-part PCR amplification, using pRB1 as a template. The upstream region was amplified using the p400cpeC primer paired with a series of different inside primers, and the downstream region was amplified with PC1Gjn and a series of different inside primers. All inside primers were designed to generate a random sequence to replace portions of the region comprising positions −150 to −90 (see Results) and to provide the overlapping sequence needed to allow annealing of the two initial amplification products in the third PCR amplification, which was conducted using the primers 400cpeC and PC1Gjn, containing Splh and BamHI sites, respectively. The final product was cut with Splh and BamHI and inserted into similarly cut pRB1 to produce the constructs listed in Table 1.

To create the two plasmids containing fusions between the cpeC upstream and cpeC′ promoter and 5′ leader regions (with and without the L box) (Table 1), the region of the cpeC′ operon from positions −421 to −71 was PCR amplified either with primers 400cpeC and cpc1Lbox1 or with primers 400cpeC and cpc1Lmut1.

A second PCR amplified the cpc1′ promoter and 5′ leader regions from positions −70 to +286, using either primers PCC3′B and cpc1Lbox2 or primers PCC3′B and cpc1Lmut2. The cpc1Lbox1 and cpc1Lbox2 primers and the cpc1Lmut1 and cpc1Lmut2 primers were designed to introduce complementary sequences into both products that were used to anneal the products from the first two amplifications, which were then reamplified using primers 400cpeC and PCC3′B. The product was cut with Splh and BamHI and inserted into similarly cut pRB1 to create the plasmids.

Bioinformatics. Nucleotide sequences for the CCA-capable cyanobacteria Pseudanabaena sp. PCC 7409 (http://www.crbip.pasteur.fr/fiches/fichecata.jsp?crbip=pageMain, respectively.

GUS assays. Beta-glucuronidase (GUS) assays were modified from previous protocols (6, 24). F. diplosiphon transformants harboring the respective promoters gusA were grown in 50 ml liquid BG-11 with 10 gusA protocols (6, 24). F. diplosiphon /H11005 speciesTag Pseudanabaena create the plasmids.

Both products that were used to anneal the products from the first two amplifications in the third PCR amplification, which was conducted using the primers 400cpeC and PC1Gjn, containing Splh and BamHI sites, respectively. The final product was cut with Splh and BamHI and inserted into similarly cut pRB1 to produce the constructs listed in Table 1.

To create the two plasmids containing fusions between the cpeC′ upstream and cpeC′ promoter and 5′ leader regions (with and without the L box) (Table 1), the region of the cpeC′ operon from positions −421 to −71 was PCR amplified either with primers 400cpeC and cpc1Lbox1 or with primers 400cpeC and cpc1Lmut1. A second PCR amplified the cpc1′ promoter and 5′ leader regions from positions −70 to +286, using either primers PCC3′B and cpc1Lbox2 or primers PCC3′B and cpc1Lmut2. The cpc1Lbox1 and cpc1Lbox2 primers and the cpc1Lmut1 and cpc1Lmut2 primers were designed to introduce complementary sequences into both products that were used to anneal the products from the first two amplifications, which were then reamplified using primers 400cpeC and PCC3′B. The product was cut with Splh and BamHI and inserted into similarly cut pRB1 to create the plasmids.

Bioinformatics. Nucleotide sequences for the CCA-capable cyanobacteria Pseudanabaena sp. PCC 7409 (http://www.crbip.pasteur.fr/fiches/fichecata.jsp?crbip=pageMain, respectively.

GUS assays. Beta-glucuronidase (GUS) assays were modified from previous protocols (6, 24). F. diplosiphon transformants harboring the respective promoters gusA were grown in 50 ml liquid BG-11 with 10 gusA protocols (6, 24). F. diplosiphon /H11005 speciesTag Pseudanabaena create the plasmids.

RESULTS

L box contexts are comparable in promoters from several CCA-capable cyanobacteria. In F. diplosiphon, the L boxes upstream of the pcyA gene and the cpc2 and cpeC operons differ in orientation and position, and these differences correspond to the expression of these genes in either RL or GL (1) (Fig. 1A). The L boxes within the pcyA and cpc2 promoters, which are both upregulated in RL, are in the same orientation and location relative to the transcription start site. We designate this the “forward” orientation of the L box relative to the direction of transcription of the affected gene. The L box upstream of the GL-upregulated cpeC operon is in the opposite direction, which we refer to as the “reverse” orientation. The pcyA and cpc2 L boxes act as activating elements in RL, while the cpeC L box acts as a repressing element in RL (1, 32). The locations, orientations, and flanking sequences of these three L boxes are all different, even though the core repeats within them are completely conserved, with the exception of a single base pair within the cpeC operon.

Thus far, the only cyanobacterium that is capable of CCA whose completely sequenced genome is available is Synechococcus sp. PCC 7335. Its genome contains the genes that encode the three components of the Rca system, RcaE, RcaF, and RcaC, as well as L boxes immediately 5′ of the pcyA gene and the cpc2 and pebAB operons (Fig. 1A). In addition, the L boxes upstream of pcyA and the cpc2 operon in Synechococcus sp. PCC 7335 are in the forward orientation, while the pebAB operon, which has been shown to be CCA regulated and more highly expressed in GL in F. diplosiphon (2), is flanked by an L box in the reverse orientation. Thus, it is possible that Synechococcus sp. PCC 7335 uses the Rca system and L boxes to control CCA in a manner quite similar to that of F. diplosiphon. Pseudanabaena sp. PCC 7409 also undergoes CCA and contains an L box upstream of the cpc2 operon, in the forward orientation and the same position relative to transcription as that in F. diplosiphon (13, 44). Collectively, these data suggest that although some features of the Rca regulatory pathway differ in these species, they share basic features of this system, including the correlation between the orientation of the L boxes and their apparent use in activating and repressing gene expression in RL.

The L box repeats and flanking sequences for F. diplosiphon, Synechococcus sp. PCC 7335, and Pseudanabaena sp. PCC 7409 are shown in Fig. 1B, and a comparative analysis of the first two is provided in Fig. 1C. There is an extremely high level of sequence conservation within the six known F. diplosiphon repeats (41 of 42 bp identical) and somewhat less, but still strong, conservation within the Synechococcus sp. PCC 7335 repeats (37 of 42 bp identical) and between the repeats of the three species (36 of 42 bp identical between F. diplosiphon and Synechococcus sp. PCC 7335 and 13 of 14 bp identical between F. diplosiphon and Pseudanabaena sp. PCC 7409). It is also noteworthy that the differences between the repeats occur only at the outer nucleotide of the 5′ or 3′ end of the repeats, never in the internal regions. The sequences flanking the repeats (defined as the 5 bp on each outer side of the repeats) and the 4 bp between the repeats are much more diverged. Only the F. diplosiphon cpc2 and cpeC flanking sequences (10 of 14 bp identical), the cpc2 flanking sequences from F. diplosiphon and Synechococcus sp. PCC 7335 (6 of 14 bp identical), and the cpc2 flanking sequences from F. diplosiphon and Pseudanabaena sp. PCC 7409 (7 of 14 bp identical) showed above-random similarity. Overall, the variations in both the L box repeat and flanking sequences appear likely to affect the activity of this cis-acting element.

Variation in sequences flanking the L box repeats controls promoter activity. We examined the relative RL activation capabilities of the cpc2, pcyA, and cpeC L boxes and the influence of the DNA sequences flanking the direct repeats of each of these by placing them into the non-CCA-regulated cpc1 promoter and 5′ leader region. Each was joined as a translational fusion to the gusA reporter gene (Fig. 2A). Every L box region included 5 bp of flanking DNA on each side, since this was the region protected by RcaC in previous footprinting studies (32), and was positioned so that the center of the 4-bp intervening sequence between the two repeats was located at position −32/−33 relative to the cpeC transcription start site.
(11) (Fig. 2B). The base construct, pPCc412G, consisting of only the cpe1 promoter and 5’ leader region, was slightly more active in GL than in RL (Fig. 2C) (6). A negative-control construct (pPCc1) containing 28 bp of random sequence in the same position as constructs containing the L boxes showed reduced activity that was also higher in GL than in RL (Fig. 2C). The cpe2 L box (pPCc3) conferred the greatest RL induction and overall activity of the three, with 8-fold higher RL activity than that of the negative control. In fact, the addition of this L box to the cpe1 promoter increased its activity to a level comparable to that measured for the cpe2 promoter and cpe1 5’ leader sequence translationally fused to gusA (Fig. 2C, pPC2-C), suggesting that this L box is the major, and likely only, component controlling the RL-mediated transcriptional activation of cpe2. This is supported by previous results from a cpe2 deletion analysis (6). The pcyA L box (pPCc11) provided significantly less RL induction than the cpe2 L box (only 4-fold above the negative-control level), although the activities of these two promoters in GL were very similar. Since the 7-bp core repeats of the pcyA and cpe2 L boxes are identical (Fig. 2B), the sequences flanking the repeats must be the major determinants of the capacity for RL transcriptional activation. The cpeC L box (pPCc10), which is a repressing element in its native promoter (32), functioned as an activator when its orientation was reversed and it was placed in the same promoter location as the cpe2 L box (Fig. 2C). However, it was less effective than the cpe2 and pcyA L boxes at transcriptional activation in RL, increasing levels to only 3 times that of pPCc1. Compared to the other two L boxes, there is a 1-bp difference within the direct repeats of the cpeC L box (Fig. 2B) that affects its capacity to activate transcription in RL, since changing only this base pair within the context of a cpe2 L box and flanking sequence (pPCc9) decreased RL transcriptional induction by approximately 50% (Fig. 2C). Therefore, the relative weakness of the cpeC L box must be due primarily to this single base pair change, although the sequences flanking the direct repeats must also contribute, since a promoter with the 1-bp mutant version of the cpe2 L box was still 1.4 times more active in RL than the promoter containing the cpeC L box element (Fig. 2C).

L box flanking sequences have a major effect on RL-mediated transcriptional activation. We analyzed the influence of the 5-bp regions upstream and downstream of the cpe2 and pcyA L boxes to further assess their importance in RL transcriptional activation. Two versions of each of these were created, one containing the normal flanking sequences and one in which these regions had been modified (Fig. 3A). These were placed into the cpe1 promoter and 5’ leader region, which was translationally joined to gusA (Fig. 3B). The cpe2 L box was positioned as in Fig. 2B, while the pcyA L box was placed 2 bp closer to the transcription start site to match its location in its native promoter (1). For the cpe2 L box, the loss of the upstream and downstream flanking sequences led to a decrease in transcriptional activity of approximately 35% in RL and 16% in GL (Fig. 3C). This shows that the cpe2 L box flanking sequences contribute significantly to the effectiveness of L boxes as transcriptional activation elements, while the pcyA flanking sequences apparently contribute little more to the effectiveness of that L box than the direct repeat sequences alone.
Capacity for RL activation depends on L box orientation. A central question concerning the function of L boxes is whether their orientation is critical for their activity, particularly for those relatively close to the transcription start site. We addressed this by reversing the orientation of the cpeC L box within the cpc1 promoter while maintaining its position relative to the transcription start site (Fig. 4A and B). This reorientation involved both the direct repeat sequence and the flanking sequence, as shown in Fig. 2B. The results clearly show that the orientation of the L boxes that activate transcription in RL is critical for their function (Fig. 4C). An L box placed in its normal orientation imparted RL responsiveness (pPCc3), while reversing the orientation of the L box actually led to repression of transcriptional activity in RL (pPCc8). This was presumably the result of the continued binding of RcaC to the L box in RL, but in a manner that interfered with DNA-dependent RNA polymerase binding and/or transcription initiation. These results suggest that very specific contacts must be made between RcaC and the basal transcription machinery to achieve enhanced transcriptional activity in RL.

FIG. 5. Effect of L box flanking sequences on the strength of their activity. (A) Diagrams representing the constructs used for panel C. The cpeC promoter and 5' leader regions were fused to the gusA reporter gene, with the L box from the cpc2 operon substituted for the cpc1 sequence centered at position –33 to –32, in the forward or reverse orientation. (B) Sequence alignment of the cpeC promoter containing the cpc2 L box in either the wild-type (pPCc3) or reverse (pPCc8) orientation. All numbering is relative to the transcription start site for the cpc1 operon. L box repeats are underlined and substituted sequences (bold) are shown in the cpc1 promoter. (C) Relative mean rates of GUS activity per mg of protein for cell lysates of F. diplosiphon transformed with the indicated plasmids and grown in RL or GL. The 100% value was 741.2 nmol of product per mg of protein per min and was derived from cells transformed with pPCc3 and grown in RL. At least four independently transformed lines were tested for each plasmid and light condition. The RL-to-GL ratio of mean GUS activities is provided in parentheses above each construct. Error bars show standard errors of the means.
An element(s) that activates high-level expression of cpeC is located upstream of the L box. To determine whether the cpeC promoter and upstream sequences contained any activating DNA elements, we constructed a series of 5′ deletions of this region. The base construct, pRB1, contained the cpeC promoter and upstream sequences from positions −421 to +286 fused to the 5′ leader region of cpc1 (from positions +10 to +286, encoding the first 11 amino acids), which had been joined translationally to gusA (Fig. 6A). This construct was expressed 2.5 times more in GL than in RL (Fig. 6B), and this difference was due to the RcaC system operating through the L box (32).

Analysis of constructs in which regions upstream of cpeC had been removed demonstrated that the sequence down to position −150 could be changed without affecting GL expression but that removal of the sequence from positions −150 to −110 resulted in the loss of high-level expression in GL and the loss of CCA regulation (Fig. 6B). Further loss of the cpeC sequence down to position −52, which eliminated the L box, resulted in slightly less overall expression than that with the −110 truncation, with a continued loss of the CCA response. Taken together, these results suggest that there is one or more positively acting element in the sequence between positions −150 and −110 upstream of the cpeC start site and that this element(s) is required for the normal CCA response of cpeC.

The cpeC activating element(s) between positions −150 and −110 is redundant or extended. Because the DNA sequence from positions −150 to −110 of cpeC increases transcriptional activity, particularly in GL (Fig. 6B), it was further examined for an activating element(s). This region contains a direct repeat, shown in Fig. 7A and B. Four internal substitution mutants were created, using pRB1. These replaced the sequence from positions −142 to −118 (pRB1ΔA), which eliminated direct repeat A (5′-CCATTACCCATT2CCATTCCCCATT A-3′); the sequence from positions −114 to −100 (pRB1ΔB), which eliminated region B, immediately downstream of direct repeat A; the sequence from positions −107 to −90...
which eliminated region C, immediately upstream of the L box; and the sequence from positions 142 to 100 (pRB1ΔC), which eliminated region C, immediately upstream of the L box; and the sequence from positions −142 to −100 (pRB1ΔAB), which combined the substituted regions of pRB1ΔA and pRB1ΔB and eliminated the A and B regions (Fig. 7A and B).

The GL expression level did not decrease for the pRB1ΔA and pRB1ΔB mutants, as would be expected if a sufficient amount of a single essential activating element had been eliminated in either of these regions (Fig. 7C). The GL expression level was somewhat elevated in the combination mutant pRB1ΔAB, which suggests that there are no elements between positions −142 and −100 that work redundantly to transcriptionally activate cpeC. For construct pRB1ΔC, the activity increased significantly in both GL and RL compared to that measured for pRB1.

**FIG. 7.** Internal substitution analysis of the region immediately upstream of the cpeC L box to examine activating elements. (A) Diagram representing the constructs used for panel C. The cpeC upstream, promoter, and 5′ leader from positions −421 to +38 were fused to the cpc1 5′ leader sequence from positions +10 to +286. A direct repeat centered at position −130 (A), a 15-bp region centered at position −107 (B), and a region upstream of the L box (C) were identified as possible positively acting elements or regions that operate upstream of the L box and were eliminated by replacement with a random sequence. (B) Sequence alignments of the various constructs used for panel C. The cpeC upstream region is shown from the L box (boxed region) at position −71 up to position −150. The bold underlined sequences were mutated to a random sequence in order to remove possible cis-acting regions A, B, and C, which are denoted at the top. Arrows denote the direction of the A direct repeat. (C) Relative mean rates of GUS activity per mg of protein for cell lysates of *F. diplosiphon* transformed with the indicated plasmids and grown in RL or GL. The mean value was 263.4 nmol of product per mg of protein per min and was derived from cells transformed with pRB1 and grown in GL. At least four independently transformed lines were tested for each plasmid and light condition. The GL-to-RL ratio of mean GUS activities is provided in parentheses above each construct. Error bars show standard errors of the means.

cpeC L box function requires cpeC sequences downstream of the L box. The results in Fig. 6 demonstrated that the L box which represses cpeC expression in RL (32) cannot function in the absence of sequences upstream of position −110, which are required for high levels of cpeC transcripational activation. It is possible that cpeC sequences downstream of the L box are also required for a sufficiently high level of cpeC expression in GL to allow CCA activity. To test this, the region of cpeC from positions −412 to −71 from pPCc412G was replaced with either the native cpeC sequence from positions −421 to −71, which contains the L box (pPCc12), or the equivalent region with a mutated L box sequence (pPCc13) (Fig. 8A). Neither of these constructs was expressed in a significantly different fashion in either RL or GL from that for the base construct pPCc412G, even though they were all highly expressed (Fig. 8B). This indicates that the CCA regulation of cpeC transcriptional activity through the repressing action of the L box cannot occur without specific DNA sequences downstream of the L box, specifically between positions −70 and +30, since only this region was different between pPCc12 and pRB1 (Fig. 6), which was CCA regulated.

**DISCUSSION**

The functional studies presented here demonstrate that the sequence context, position, and orientation of L boxes are all important for their function as CCA activating and repressing elements during growth in RL, a hypothesis that emerged from
FIG. 8. Effectiveness of the cpeC L box in the absence of downstream cpeC sequence. (A) Diagrams representing the constructs used for panel B. The cpc1 upstream, promoter, and 5′ leader regions (pPCc412G), the cpeC upstream region from positions −400 to −71 (with the L box) fused to the cpc1 promoter and 5′ leader sequence from positions −70 to +286 (pPCc12), or the cpeC upstream region from positions −400 to −71 and containing random sequence substituted for the L box fused to the cpc1 promoter and 5′ leader sequence from positions −70 to +286 (pPCc13) were used in translational fusions to drive the expression of gusA. (B) Relative mean rates of GUS activity per mg of protein for cell lysates of F. diplosiphon transformed with the indicated plasmids and grown in RL or GL. The mean value was 428.6 nmol of product per mg of protein per min and was derived from cells transformed with pPCc412G and grown in GL. At least four independently transformed lines were tested for each plasmid and light condition. The GL-to-RL ratio of mean GUS activities is provided in parentheses above each construct. Error bars show standard errors of the means.

This might be partly because the 4 bp that differ between the cpe2 and cpeC L box flanking sequences are critical for high-affinity RcaC binding. In addition, the single nucleotide change at the end of one of the cpc1 repeats (compared to the cpe2 repeats) (Fig. 2B) also influences the strength of RcaC binding, as demonstrated by the 2-fold lower activity of the cpe2 L box containing this change than that of the wild-type L box (Fig. 2C).

The L box repeat sequences are also quite highly related between species. Despite the fact that neither Synechococcus sp. PCC 7335 nor Pseudanabaena sp. PCC 7409 is closely related to F. diplosiphon within the cyanobacteria (23), the repeats present in the cpe2 promoters of these three species are identical in at least 12/14 positions (Fig. 1D). Overall, the L box repeats appear to be more highly conserved between promoters than the repeats found within the binding sites of other members of this response regulator class, such as OmpR and PhoB (5, 19, 31), although it remains possible that many other, less highly conserved RcaC binding sites exist but have not yet been identified. This L box conservation may be evidence of recent lateral gene transfer of the Rca system between cyanobacterial species or of a strong fitness advantage that is conferred by the level of interaction that exists between RcaC and the consensus repeat sequences within the L box.

We note that in the experiments presented here, the 10-fold difference in RNA abundance that is normally measured for the cpeC operon during CCA (2, 4, 40) was reduced to a 2- to 3-fold response for pRB1, which consists of the cpeC sequence from positions +38 to −421 and the cpef 5′ leader region from positions +10 to +286 (Fig. 6). This discrepancy between GUS activity differences and cpeC RNA accumulation differences has been observed previously (32) and is the result of the absence of influence of the second CCA regulatory pathway, called the Cgi system (30), which appears to function posttranscriptionally through the 5′ leader of the cpeC operon (our unpublished data).

For each of the L boxes identified thus far (Fig. 1C), the fact that the repeats are either identical or differ in only one of the seven positions may provide a clue to how RcaC interacts with L boxes. PhoB and OmpR, the best-characterized members of this group of response regulators, operate through different mechanisms. The dimerization and subsequent DNA binding enhancement of PhoB appear to be driven by its phosphorylation, while OmpR appears to initially bind to DNA as a monomer and then to undergo enhanced phosphorylation, dimerization, and binding of a second monomer (3, 35, 39). The sequence similarity between the half-sites within the DNA target sequence is quite high for PhoB (5), while for OmpR the greater differences in the sequences of the half-sites result in different binding affinities and allow the stepwise addition of OmpR monomers to the DNA (39). Therefore, the high degree of sequence conservation within the repeat sequences of the L boxes, together with the fact that RcaC appears to be a dimer in its unbound, active state (L. Li and D. M. Kehoe, unpublished data), suggests that the mechanism underlying RcaC-L box interactions may be more similar to that controlling PhoB than to that controlling OmpR.

The activating function of the L box is clearly orientation specific (Fig. 4C). This differs from the apparent situation for RapB, another cyanobacterial OmpR/PhoB class factor which...
binds direct repeat HLR1 elements within the promoters of light-intensity-regulated genes and, depending on the gene, can act as either an activator or a repressor under low-light conditions (14, 25, 26, 41, 42). Although the dual function of RpaB via HLR1 elements is similar to the action of RcaC in RL through the L box, there are indications that the mechanisms through which they act are quite different. The most notable of these is that unlike L boxes, both positively and negatively acting HLR1 elements can be present near or just upstream of the transcription start sites of the genes they regulate. More importantly, although not yet tested directly, HLR1 elements also appear to be capable of functioning in either orientation, regardless of whether they are acting positively or negatively. For CCA regulation, the inability of an either orientation, regardless of whether they are acting positively or negatively. 

In summary, the data presented here advance our understanding of the mechanisms of action underlying light regulation and OmpR/PhoB class transcription factors in cyanobacteria. This study has further characterized the features of the L box, which is the binding site of RcaC, the best-studied cyanobacterial OmpR/PhoB family member. Our results demonstrate that the L box location, orientation, and both core repeat and flanking sequences all contribute in very specific ways to the strength of this element and its ability to function as an activator or repressor.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (grant MCB-0519433 to D.M.K.).

We thank Matt Carscely, Peter Gracheck, and Noyan Olcum for their assistance with the work described here. We also thank the other members of the Keohoe laboratory for thoughtful discussions and comments on this work.

REFERENCES


