Occurrence of Cyclic di-GMP-Modulating Output Domains in Cyanobacteria: an Illuminating Perspective

Marco Agostoni,a,b Benjamin J. Koestler,c Christopher M. Waters,a,c Barry L. Williams,d Beronda L. Montgomerya,b,e

ABSTRACT Microorganisms use a variety of metabolites to respond to external stimuli, including second messengers that amplify primary signals and elicit biochemical changes in a cell. Levels of the second messenger cyclic dimeric GMP (c-di-GMP) are regulated by a variety of environmental stimuli and play a critical role in regulating cellular processes such as biofilm formation and cellular motility. Cyclic di-GMP signaling systems have been largely characterized in pathogenic bacteria; however, proteins that can impact the synthesis or degradation of c-di-GMP are prominent in cyanobacterial species and yet remain largely underexplored. In cyanobacteria, many putative c-di-GMP synthesis or degradation domains are found in genes that also harbor light-responsive signal input domains, suggesting that light is an important signal for altering c-di-GMP homeostasis. Indeed, c-di-GMP-associated domains are often the second most common output domain in photoreceptors—outnumbered only by a histidine kinase output domain. Cyanobacteria differ from other bacteria regarding the number and types of photoreceptor domains associated with c-di-GMP domains. Due to the widespread distribution of c-di-GMP domains in cyanobacteria, we investigated the evolutionary origin of a subset of genes. Phylogenetic analyses showed that c-di-GMP signaling systems were present early in cyanobacteria and c-di-GMP genes were both vertically and horizontally inherited during their evolution. Finally, we compared intracellular levels of c-di-GMP in two cyanobacterial species under different light qualities, confirming that light is an important factor for regulating this second messenger in vivo.

IMPORTANCE This study shows that many proteins containing cyclic dimeric GMP (c-di-GMP)-regulatory domains in cyanobacteria are associated with photoreceptor domains. Although the functional roles of c-di-GMP domain-containing proteins in cyanobacteria are only beginning to emerge, the abundance of these multidomain proteins in cyanobacteria that occupy diverse habitats ranging from freshwater to marine to soil environments suggests an important role for the regulation of c-di-GMP in these organisms. Indeed, we showed that light distinctly regulates c-di-GMP levels in *Fremyella diplosiphon* and *Synechocystis* sp. strain PCC6803. Our findings are consistent with the occurrence of c-di-GMP domains based on evolutionary origin and as an adaptation to specific habitat characteristics. Phylogenetic analyses of these domains clearly separate two distinctive clades, one composed of domains belonging predominantly to cyanobacteria and the other belonging to a mix of cyanobacteria and other bacteria. We further demonstrate that in cyanobacteria the acquisition of c-di-GMP signaling domains occurred both vertically and horizontally.
PDE activity but completely degrade c-di-GMP into two GMP molecules (8). In bacteria, the intracellular concentrations of the second messenger c-di-GMP are regulated in response to a variety of environmental stimuli. Cyclic di-GMP plays critical roles in regulating numerous cellular processes in various bacteria, including transcription, RNA turnover, biofilm formation, protein synthesis, motility, virulence, bacterial predation, and altering activities of proteins or protein complexes (9–14).

To date, the regulation of GGDEF and EAL domain-containing proteins has been studied in a limited number of cyanobacterial systems (9, 15, 16). Cyanobacteria are one of the most abundant photosynthetic organisms in aquatic environments; they are able to fix both carbon and nitrogen under aerobic conditions and thus play a key role in regulating global carbon and nitrogen cycles. The phylum Cyanobacteria represents a highly diverse group of Gram-negative bacteria that are unicellular, filamentous, or colonial. They are present in the majority of biotopes on Earth, being found in marine, freshwater, and soil environments. They have also colonized deserts (17), polar waters (18), and geothermal environments (19). They are the progenitors of chloroplasts in eukaryotic photosynthetic organisms (20), and since they are the only bacteria capable of oxygentic photosynthesis, cyanobacteria also played a critical role in the initial rise of oxygen in the atmosphere around 2.3 billion years ago (21).

Recent studies have provided evidence that light can regulate c-di-GMP levels in cyanobacteria (16, 22). Light is an important stimulus for both nonphotosynthetic and photosynthetic microorganisms, where light quality, intensity, and duration are used to control a variety of cellular processes. Light is sensed by specialized photosensory proteins that are activated by bound organic cofactors or chromophores, thereby transmitting photobiological signals to a downstream output domain (23–25). In bacteria, light-responsive domains are often linked to an output domain that impacts c-di-GMP homeostasis, thereby impacting bacterial physiology (26). These proteins likely function similar to the blue-light-sensing PDE from Klebsiella pneumoniae (26, 27); i.e., light induces a conformational change in the photosensory domain that results in alteration in the linked GGDEF, EAL, and/or HD-GYP domain that impacts c-di-GMP modulating activity. In this study, we assessed the occurrence of c-di-GMP-modulating domain-containing proteins in sequenced genomes of cyanobacteria that occupy diverse environmental habitats. We examined the abundance of proteins containing c-di-GMP-modulating domains in association with photoreceptor domains, and we propose that the widespread occurrence of these proteins indicates an important role in the environmental regulation of c-di-GMP levels in cyanobacteria. In addition, we document examples of the acquisition of new c-di-GMP signaling proteins by horizontal gene transfer (HGT) of c-di-GMP domain-containing genes and discuss which factors could have led to these events. Additionally, we quantified intracellular concentrations of c-di-GMP in Fremyella diplosiphon and Synechocystis sp. strain PCC6803 (referred to herein as Synechocystis) under white, blue, red, and green light. These two species showed different concentrations of c-di-GMP under different light qualities, which corresponded to known impacts of light on their photoadaptation in natural contexts. These studies validate the importance of light in regulating c-di-GMP homeostasis in cyanobacteria.

RESULTS

Light-responsive domains associated with c-di-GMP in cyanobacteria. Analyses of sequenced cyanobacterial genomes showed that 20 out of 37 cyanobacteria (54%) in CyanoBase possessed proteins containing c-di-GMP regulatory domains (Table 1; also, see Table S1 in the supplemental material). Putative light-responsive domains were common in cyanobacterial DGCs and PDEs (Table 1 and Fig. 1; also, see Table S1) (28); of the 398 c-di-GMP-containing proteins, 131 (33%) were associated with light-responsive domains (Table 1). A large number of putative photoreceptor genes encode GGDEF protein containing domains. Notably, only two proteins (6%) with an EAL-only domain (i.e., a protein with only the EAL domain without a GGDEF domain, as opposed to a hybrid protein that contains both GGDEF and EAL domains) were associated with photoreceptor domains. Likewise, nine HD-GYP domains (3%) were associated with photoreceptor domains (Table 1).

The number of c-di-GMP domains in cyanobacteria could be dependent on genome size. However, an alternate explanation is that habitat characteristics could have influenced the number of c-di-GMP domains in cyanobacteria. To gain insight into the likelihood of these two possibilities, we compared the size of the genome and the number of genes with the total number of c-di-GMP-related domains. Although the size of genome and the total number of genes were positively correlated with the number of c-di-GMP domains (Fig. 2A and B), these parameters explained only 43% and 61% of the total variance, respectively. These correlations indicate that the number of c-di-GMP domains are not simply correlated with genome size but may also be determined by bacterial adaptation. For instance, Trichodesmium erythraeum IMS101 and Microcystis aeruginosa NIES-843 have relatively large genomes (Table 1) with only five and three c-di-GMP domains, respectively. In contrast, the genome sizes of Synechocystis and Cyanothecae sp. ATCC 51142 are half those of T. erythraeum and M. aeruginosa genomes (Table 1) but contain 29 and 35 c-di-GMP domains, respectively. T. erythraeum inhabits tropical and subtropical oceans known to be nutrient-poor and relatively stable waters (29), and M. aeruginosa inhabits low-nutrient lakes (30). On the other hand, Synechocystis is adapted to dynamic environments—e.g., it can grow heterotrophically in the dark (31), and its metabolism is controlled largely by numerous circadian clock genes (32). Also, Cyanothecae sp. strain ATCC 51142 was isolated from intertidal waters in Texas that are subject to a range of harsh conditions (33). These examples strongly suggest that c-di-GMP domains are impacted by bacterial adaptation and that this can occur independent of genome size.

The total number of c-di-GMP domains associated with photoreceptor domains in cyanobacteria (Table 1 and Fig. 2C) is in contrast with the total number of c-di-GMP-domain-containing proteins associated with photoreceptors from all bacteria. In other bacterial genomes, the most common photoreceptor domain associated with c-di-GMP domains is the BLUF (sensor of blue light using FAD) domain (40%), followed by the LOV (light-oxygen-voltage) domain, a subfamily of the PAS domain family (21%), and the phytochrome-characteristic GAF domain (8%) (34), used by cyanobacteria to sense UV-A, blue, green, and/or red light (35–38). Our analyses showed different results in that c-di-GMP domains in cyanobacteria were mostly associated with the GAF domain (58%) and the PYP domain of xanthopsin photoreceptors.
and were found less frequently in association with the LOV domain (6%) (Table 1). A number of LOV domain-containing proteins from cyanobacteria have already been shown to be associated with regulation of c-di-GMP levels or with c-di-GMP signaling (22, 39, 40). We did not identify any BLUF domain associated with DGCs or PDEs in the available cyanobacterial genomes; however, this finding is not unexpected based on the low abundance of BLUF domains in cyanobacterial genomes (41).

Some DGC proteins possess an allosteric inhibition site (I site) proximal to the GGDEF active site (42). This c-di-GMP-dependent allosteric site, characterized by an RXXD motif, is important for controlling DGC activity; when levels of c-di-GMP are high, the second messenger can bind the RXXD motif, thereby repressing the DGC activity. Among all the GGDEFs associated with photoreceptor domains in cyanobacteria, the allosteric site was found in 94% of GGDEF-only proteins and 81% of hybrid proteins (see Table S2 in the supplemental material). These values are slightly higher than those for all the GGDEFs that are not associated with photoreceptor domains in cyanobacteria, where the I site was found in 82% of GGDEF-only proteins and 63% of hybrid proteins (see Table S2). The frequency of the RXXD motif in cyanobacteria is different from that in other bacteria, where only half of the DGCs possessed an allosteric site (43). The high occurrence of GGDEF proteins containing an I site suggests that cyanobacteria tightly control c-di-GMP synthesis when DGCs are associated with photoreceptors.

Cyclic di-GMP domains: vertical and horizontal gene transfer. Cyclic di-GMP domains are widespread in bacteria; thus, it

![Table 1: Numbers of c-di-GMP domain-containing proteins and subsets containing photoreceptors domains in sequenced cyanobacterial genomes](https://example.com/table1)

**Table 1** Numbers of c-di-GMP domain-containing proteins and subsets containing photoreceptors domains in sequenced cyanobacterial genomes

| Strain and genome size (bp) | No. of c-di-GMP domains<sup>a</sup> | No. of photoreceptors associated with c-di-GMP |
|-----------------------------|-----------------------------------|---------------------------------------------|
|                             | GGDEF | GGDEF and EAL | EAL | HD-GYP | Total | GAF | PYP | LOV |
| Acaryochloris marina MBIC11017; 6,503,724 | 28 (7) | 26 (23) | 6 (0) | 0 (0) | 60 (13) | 8 | 0 | 0 |
| Arthrospira platensis NIES-39; 6,788,435 | 21 (38) | 13 (38) | 4 (0) | 0 (0) | 38 (34) | 7 | 2 | 4 |
| Cyan methane sp. PCC 7424; 5,942,652 | 19 (26) | 14 (36) | 0 (0) | 2 (0) | 35 (29) | 6 | 3 | 1 |
| Cyan methane sp. ATCC 51142; 4,934,271 | 12 (50) | 18 (56) | 3 (33) | 1 (0) | 34 (50) | 3 | 13 | 1 |
| Synechocystis sp. PCC 6803; 3,573,471 | 13 (46) | 9 (33) | 5 (0) | 2 (0) | 29 (31) | 7 | 1 | 1 |
| Cyan methane sp. PCC 8801; 4,679,413 | 13 (69) | 10 (20) | 2 (50) | 1 (0) | 26 (46) | 10 | 2 | 0 |
| Nostoc punctiforme ATCC 29133; 8,234,322 | 11 (55) | 10 (70) | 1 (0) | 2 (0) | 24 (54) | 2 | 10 | 1 |
| Cyan methane sp. PCC 7425; 5,374,574 | 7 (14) | 12 (42) | 2 (0) | 1 (100) | 22 (32) | 5 | 2 | 0 |
| Synechococcus elongatus PCC 6301; 2,696,255 | 9 (56) | 8 (63) | 1 (0) | 2 (100) | 20 (60) | 8 | 2 | 2 |
| Synechococcus elongatus PCC 7942; 2,695,903 | 9 (56) | 8 (63) | 1 (0) | 2 (100) | 20 (60) | 8 | 1 | 3 |
| Anaabaena variabilis ATCC 29413; 6,365,727 | 8 (0) | 6 (50) | 2 (0) | 2 (0) | 18 (18) | 2 | 0 | 1 |
| Anabaena sp. PCC 7120; 6,413,771 | 8 (13) | 6 (33) | 1 (0) | 2 (0) | 17 (17) | 2 | 0 | 1 |
| Synechococcus sp. PCC 7002; 3,008,047 | 7 (14) | 6 (33) | 2 (0) | 0 (0) | 15 (20) | 0 | 2 | 1 |
| Thermosynechococcus elongatus BP-1; 2,593,857 | 5 (20) | 5 (60) | 1 (0) | 1 (100) | 12 (42) | 5 | 0 | 0 |
| Synechococcus sp. IA-2-3Bp(a2-13); 3,046,682 | 4 (25) | 3 (20) | 0 (0) | 3 (33) | 7 (29) | 2 | 0 | 0 |
| Synechococcus sp. IA-3-3Ab; 2,932,766 | 3 (0) | 0 (0) | 1 (0) | 2 (50) | 6 (17) | 1 | 0 | 0 |
| Synechococcus sp. C9311; 2,606,748 | 1 (0) | 3 (0) | 1 (0) | 0 (0) | 5 (0) | 0 | 0 | 0 |
| Trichodesmium erythraeum IMS101; 7,500,106 | 1 (0) | 3 (0) | 0 (0) | 1 (100) | 5 (20) | 1 | 0 | 0 |
| Microcystis aeruginosa NIES-843; 5,842,795 | 1 (0) | 1 (0) | 0 (0) | 1 (0) | 3 (0) | 0 | 0 | 0 |
| Gloeobacter violaceus PCC 7421; 4,659,019 | 1 (0) | 0 (0) | 0 (0) | 1 (0) | 2 (0) | 0 | 0 | 0 |
| Total | 181 (31) | 158 (40) | 33 (6) | 26 (35) | 398 (33) | 77 | 38 | 16 |

<sup>a</sup> Values in parentheses are percentages of c-di-GMP modulating domains associated with photoreceptors.
has been suggested that c-di-GMP is an ancient second messenger (44). However, an alternate explanation, which is not mutually exclusive, is that cyanobacteria acquired c-di-GMP-associated domains through HGT events. We examined whether the presence of c-di-GMP domains was a result of vertical transfer events or if HGT events regularly occurred in cyanobacteria. Genes can be considered acquired via HGT events when phylogenetic analysis of genes shows clustering of distantly related species. We performed this analysis for the conserved EAL-only proteins by generating a phylogenetic tree of 15 noncyanobacterial EAL domain proteins and 24 conserved cyanobacterial EAL-only proteins.

The phylogenetic analysis of multiple conserved EAL domain sequences showed two distinctive clades (Fig. 3; also, see Fig. S1 in the supplemental material). One clade was composed of only cyanobacteria EAL-only domains, with the exception of a single noncyanobacterial sequence, and the other was composed of a mix of noncyanobacterial and cyanobacterial EAL-only domains. The only species having genes widely distributed in both clades was *Acaryochloris marina*, a unique cyanobacterium that uses a distinctive light acclimation method based on chlorophyll d, a far-red- and infrared-absorbing chlorophyll (45). It would be interesting to explore why this species possesses such a great diversification of EAL domains. In addition, the gene *A28LD_0392* from *Idiomarina* sp. strain A28L, a Gram-negative, aerobic, flagellate gammaproteobacterium present in a wide range of aquatic saline habitats (46), was the only noncyanobacterial gene clustered within the cyanobacterium-predominant clade. This gene exhibited high amino acid sequence identity to *Synechocystis slr6110* (E value, 10^-110). Notably, the *slr6110* gene is not incorporated in the genomic DNA but found on the *Synechocystis* plasmid pSYSX and itself could have been acquired via an HGT event. When we compared the sequences of *slr6110* and *A28LD_0392*, we observed two long conserved regions and one small deletion not present in the other genes, suggesting a shared evolutionary history unique to these two sequences (Fig. 4).

After performing the phylogenetic analyses, we searched for rare genetic events like indels (insertions and deletions) in our alignment that could be used to link closely related genes despite disparate evolutionary histories for the remainder of the genome, indicative of HGT events. The EAL domains encoded by *cya_1130*, *cocor_05649*, *mxam_2424*, and *staur_3026* exhibited a conserved deletion of 5 amino acids (Fig. 4). The genes *cocor_05649*, *mxam_2424*, and *staur_3026* are from a group of bacteria called myxobacteria and were closely related to the cyanobacterial gene *cya_1130*, with an E value of <10^-78. However, the genomes of these four species possessed very similar G+C content values, constraining the use of G+C content analysis between the four species. Also, the cyanobacterial gene *pcc8801_0177* and the gene *pstab_3153* from *Pseudomonas stutzeri* possessed insertions and deletions of amino acids that were not present in the other EAL
domains (Fig. 4). The presence of these indels led us to hypothesize that the EAL domains from the proteins encoded by cya_1130, cocor_05649, mxam_2424, and staur_3026 and those from the proteins encoded by pcc8801_0177 and pstab_3153 are closely related based on independent phylogenetic analyses (Fig. 3; also, see Fig. S1 in the supplemental material), which exclude sequence alignment gaps. In addition, pcc8801_0177 had a G/H content of 53.1%, while the remainder of the Cyanothece sp. PCC 8801 genome had a G/H content of 39.8%, suggesting that pcc8801_0177 was also acquired via another HGT event. These rare genetic events indeed confirmed the predictions based on the phylogenetic analyses that those EAL domains were acquired through HGT. Thus, our results suggest that a c-di-GMP signaling system evolved early in the phylum Cyanobacteria, indicating an ancient origin of these signaling systems, but that HGT events enriched the number of these domains in cyanobacterial genomes.

**Comparison of the intracellular levels of c-di-GMP between cyanobacterial species under different light qualities.** In cyanobacteria, more than 30% of c-di-GMP domains are associated with photoreceptor domains (Fig. 2C and Table 1). This finding suggests that sensing of light by photoreceptor domains may impact cellular c-di-GMP levels through modulation of enzymatically active DGCs and PDEs. Therefore, we quantified the levels of intracellular c-di-GMP of two cyanobacterial species, the filamentous freshwater cyanobacterium *F. diplosiphon* and the single-cell model cyanobacterium *Synechocystis*, exposed to different light qualities. These two species are adapted to specific environmental niches, and we hypothesized that these cyanobacteria would differently regulate intracellular levels of c-di-GMP under diverse light conditions to which these organisms are adapted in natural contexts. Notably, *F. diplosiphon*, a cyanobacterium able to change its pigmentation to maximally absorb available red or green wavelengths (47), showed higher intracellular levels of c-di-GMP under white light or red light than under green light or blue light ($P < 0.05$) (Fig. 5A). *Synechocystis* inhabits freshwater lakes and has been intensively studied for its capability to sense and respond to blue light. For instance, *Synechocystis* can grow heterotrophically in the dark but requires blue light for a few minutes each day (31). *Synechocystis* dissipates excess captured light energy through a blue-light-dependent activation of a soluble carotenoid binding protein, i.e., orange carotenoid protein, that when associated with phycobilisomes results in photoprotective fluorescence quenching (48). Finally, blue light exposure has been shown to have impacts on *Synechocystis* motility. One study provided evidence for blue-light dependent accumulation of cyclic AMP (cAMP) that was associated with cellular motility (49). In a separate study, *Synechocystis* was demonstrated to exhibit an inhibition of phototaxis by blue light that was predicted to be associated with blue light-dependent c-di-GMP accumulation (16). Though disparate movement responses were observed for wild type in these two studies, phototaxis is apparently controlled by blue light-dependent regulation of the accumulation of second messengers cAMP (49) and c-di-GMP (16). *Synechocystis* showed higher levels of c-di-GMP under blue light ($P < 0.05$) than under white light, green light, or red light. Red light exposure resulted in lower c-di-GMP levels than did green light ($P < 0.05$) (Fig. 5B). Notably, the intracellular c-di-GMP level was lower in *Synechocystis* than in *F. diplosiphon* under all light except blue light. These results suggest that different cyanobacteria respond differently in regulating c-di-GMP levels in response to external light conditions, likely due to adaptation to specific environmental niches.
Occurrence of c-di-GMP domains. Light sensing is critical for cyanobacteria to respond to different light spectra. The high number of GGDEF-only domains associated with photoreceptor domains indicates that light predominately controls synthesis rather than degradation of c-di-GMP. Cyanobacteria differ from other bacteria regarding the number of photoreceptor domains associated with c-di-GMP domains. We demonstrated that many proteins containing c-di-GMP-regulatory domains were found to be associated with blue- and red-light-dependent photoreceptor domains. The most common photoreceptor associated with c-di-GMP in noncyanobacteria was the BLUF domain (41), whereas in cyanobacteria, such associations in the genomes of sequenced cyanobacteria were rare. Instead, in cyanobacteria, the GAF domain in association with c-di-GMP domains was more common. These results are in accordance with the general distribution of the photoreceptor types among the bacterial groups (41). Although the majority of cyanobacteria contain more GAF domains than PYP domains associated with c-di-GMP genes, Synechococcus sp. strain PCC 7002, which inhabits mud areas, Nostoc punctiforme ATCC 29133, which inhabits soil, and Cyanothece sp. strain ATCC 51142, which was isolated from a warm intertidal area (50–52), possess more PYP domains than GAF domains associated with c-di-GMP genes. Since Nostoc sp. and Cyanothece sp. were found in microbial mats in their natural habitats (53, 54), we infer that the PYP domains functionally linked to c-di-GMP domains could control biofilm formation in cyanobacteria, as suggested in Gomelsky and Hoff (26), whereas the GAF domain could be an important sensor for photosensory behavior (16). This phenomenon, known as c-di-GMP signaling specificity, has been observed in other bacterial systems (55). A GAF domain associated with c-di-GMP-modulating domains has been shown to be involved in photosensory behavior in Synechocystis in response to blue light (16). The production of biofilms in cyanobacteria is a mechanism used to...
withstand harsh conditions (56), and the PYP blue light sensor may be an important receptor to regulate biofilm formation for species that are exposed more frequently to high irradiance.

**Evolutionary history of c-di-GMP in cyanobacteria.** To understand the evolutionary history of c-di-GMP in cyanobacteria, we decided to explore whether there was a correlation with organisms from a specific habitat and the presence of c-di-GMP domains. To address this question, we decided to compare the habitat characteristics of Prochlorococcus and Synechococcus strains lacking c-di-GMP domains with those of related strains that possessed them (Table 1). The presence of domains associated with c-di-GMP homeostasis in these two genera correlated with the presence of cAMP receptor proteins in cyanobacteria (57). The only species lacking genes for cAMP receptors inhabited marine oligotrophic environments, whereas species that possess cAMP receptor inhabited both marine and fresh water (57). Notably, the only genera lacking cAMP receptors were the picocyanobacteria Prochlorococcus and Synechococcus (with the exception of the species Gloeobacter violaceus). It has been suggested that organisms such as Prochlorococcus that have adapted to stable habitats have lost cAMP receptor proteins (57). Prochlorococcus, therefore, lost flexibility due to a lack of selection to maintain the receptors or as an adaptation in exchange for efficiency under stable conditions (58). In addition to cAMP receptor proteins (57), our analyses suggest that species growing under frequent environmental fluctuations cope with habitat changes by using second messengers to respond to external signals. Synechococcus strains containing c-di-GMP-modulating domains inhabit both marine and freshwater habitats and are found in nutrient-rich (eutrophic) waters (see Table S3 in the supplemental material). Prochlorococcus and Synechococcus strains lacking c-di-GMP-regulatory domains inhabit marine low-nutrient (oligotrophic) habitats (50), with the exception of CC9902, which is found in a coastal habitat (see Table S3). Currently there are no sequenced freshwater Synechococcus strains that lack these domains. Hence, the presence of c-di-GMP signaling systems may reflect the environmental characteristics of the habitats of picocyanobacteria Prochlorococcus and Synechococcus.

We also addressed the evolutionary origin of c-di-GMP signaling in cyanobacteria. G. violaceus and Thermosynechococcus elongatus BP-1 are considered the most divergent species among extant cyanobacteria (59, 60). G. violaceus possesses two degenerate c-di-GMP-modulating domains, and T. elongatus has 11 domains predicted to be associated with c-di-GMP homeostasis or signaling, half of which are associated with GAF domains (Table 1). Because the synthesis and degradation of c-di-GMP are a widespread phenomenon in cyanobacteria (Table 1) (4), it is tempting to speculate that a c-di-GMP signaling system with light-controlled c-di-GMP homeostasis evolved early in the primordial cyanobacteria. Alternatively, it is also possible that these domains were spread via HGT, although multiple HGT events from one or few donors are less likely. We performed phylogenetic analyses of conserved EAL-only domains. The phylogenetic analyses of multiple conserved EAL domain sequences identified two clades of bacteria containing EAL-only domains, one clade predominantly containing cyanobacterial c-di-GMP domains and one containing both noncyanobacterial and cyanobacterial c-di-GMP domains. The structure of these phylogenetic trees suggests that in one clade, the domains evolved vertically, whereas many of the EAL domains from the other clade likely arose as a result of HGT events within the phylum Cyanobacteria.

Although the functional roles of GGDEF, EAL, and HD-GYP proteins in cyanobacteria are only beginning to emerge, we believe that the abundance of these classes of proteins in a range of cyanobacteria that occupy diverse environmental habitats supports an important role for these proteins and the regulation of c-di-GMP levels in cyanobacteria. We also provide evidence that different cyanobacterial species possess different levels of intracellular c-di-GMP under different light qualities (Fig. 5). It is not surprising to see high levels of c-di-GMP in *F. diplosiphon*, as this species belongs to the order Nostocales. Species belonging to the order Nostocales are associated with thymolytic microbial mats (61) and can produce large amounts of biofilm, a cellular process regulated by c-di-GMP. Since *F. diplosiphon* evolved to sense red and green light by changing its pigmentation and morphology to absorb available wavelengths in the prevailing light quality at different depths in the water column (47), c-di-GMP may play a role in regulating cell morphology and buoyancy. On the other hand, in *Synechocystis*, blue light can contribute to several important processes, including heterotrophic growth, photoprotection, and phototaxis. This species showed high levels of c-di-GMP under blue light (Fig. 5B). Our results are in congruence with those reported by Savakis et al. (16), who suggested that *Synechocystis* had higher c-di-GMP levels under blue light. Indeed, the DGC Cph2 has been studied for its involvement in inhibiting phototaxis toward blue light (16). Biofilm formation in *Synechocystis* is a less common process that does not usually occur under laboratory conditions. Indeed, Schatz et al. suggested that for *Synechococcus elongatus* PCC 7942, biofilm development is self-repressed (62).

In conclusion, the regulation of c-di-GMP and light sensing may offer niche differentiation, reduce competition, and allow a variety of phytoplankton to move using pilus-based motility. Defining genes and phenotypes controlled by c-di-GMP, elucidating the mechanisms for regulating the intracellular concentration of c-di-GMP in the cell, and characterizing the regulatory pathways impacting c-di-GMP levels may help us to better understand light signaling networks in cyanobacteria.

**MATERIALS AND METHODS**

**c-di-GMP domain-containing proteins.** For the identification of GGDEF, EAL, and HD-GYP domains, we searched the Cyanobase database (http://genome.kazusa.or.jp/cyanobase), which contains genome sequences of cultured cyanobacteria (63). Conserved regions were identified manually by assessing the presence of amino acid residues known to be critical for activity in the GGDEF domain (64, 65), seven in the EAL domain (66), and 11 in the HD-GYP domain (67). The RXXD allosteric site, which is a motif found 5 to 12 amino acids before the GGDEF sequence in DGC proteins, was identified manually.

**Light-dependent receptors.** Proteins containing blue- and red-light-dependent signaling domains associated with c-di-GMP-regulatory domains were identified using gene information from Cyanobase. This option was not available for the LOV domains. Thus, to assess the potentiality of photosensory activity, LOV domains were individually screened to discriminate them from similar PAS domains with different biochemical activities. The chromophore-binding cysteine residue and the presence of 20 conserved amino acids were verified in order to identify functionally conserved LOV domains (68).

**Genome comparisons.** Phylogenetic analyses of multiple conserved EAL domain sequences were performed using TOPALi v 2.5 (69). Multiple alignments of amino acid were generated using MUSCLE. BLASTP similarity searches were used to identify a sample of homologs to produce a tree composed by cyanobacterial and noncyanobacterial EAL domains. EAL domains were selected as homologs that possess either an E value of
less than 10−80 or >50% amino acid identity. In addition, because we were focused on general patterns of evolution rather than an exhaustive description of all EAL domain evolution, we limited our taxon sampling to three homologs from each cyanobacterial species for further analyses. Phylogenetic trees were inferred using Bayesian inference in the software package MrBayes, using the Whelan and Goldman model (70) with gamma-distributed rates for the EAL domains, and the Hasegawa, Kishino, and Yano model (71) with gamma-distributed rates for 16S rRNA sequence data. MrBayes was run with two independent runs of 1,000,000 generations of three heated chains and one cold chain, with a burn-in of 25% of trees that were sampled every 10 generations. Convergence was assessed as maximum potential scale reduction factor (PSRF) values of 1.001 (EAL tree) and 1.002 (16S tree). Posterior probabilities which range from 0 to 1 were used to evaluate the statistical confidence of a particular cluster of sequences. The likelihood log was −11,774.17 for the 16S tree and −5,974.33 for the EAL-only tree.

**Cyclic di-GMP quantification and culture conditions.** *F. diplosiphon* strain SF33, a shortened-filament mutant strain that displays wild-type pigmentmentation (72), and *Synechocystis* were maintained axenically and grown in BG-11 (73) containing 20 mM HEPES at pH 8.0. Liquid cultures were adapted to fluorescent white light (Philips F32T8/TLS701/ALTO) at 35 μmol photons m−2 s−1 in glass flasks and maintained at optical densities at 750 nm (OD750) and 730 nm (OD230) of 0.6 for *F. diplosiphon* and *Synechocystis*, respectively. Flasks were maintained with shaking at 150 rpm. Cells adapted to white light were then transferred in glass flasks to white, red (λmax at 600 nm; 2506RD; LED Wholesalers), green (λmax at 530 nm; Geneva Scientific LLC), and blue light (λmax at 452 nm; 2506BU; LED Wholesalers) at 35 μmol photons m−2 s−1. Since growth rates were different under different light qualities for *Synechocystis*, flasks were inoculated at an OD230 of 0.1 for cells transferred to white light and red light, 0.4 for flasks under blue light, and 0.3 for flasks under green light. For *F. diplosiphon*, white-light, green-light, and red-light flasks were inoculated at an OD230 of 0.1 and blue-light flasks at an OD230 of 0.4. Protein and c-di-GMP quantification was performed after 5 days at OD230 of 0.6 ± 0.1.

Cyclic di-GMP levels were quantified as described in reference 55. In brief, 1.5 ml of cells were centrifuged and resuspended in 150 μl of ice-cold extraction solvent containing 40% acetonitrile, 40% methanol, and 0.1 N formic acid and mixed for 30 s by vortexing, followed by incubation for 30 min at −20°C. The cell suspension was then centrifuged at 17,000 × g for 5 min at 4°C. The supernatant was transferred to a new 1.5-ml tube and stored at −80°C until analysis. Before quantification, the supernatants were concentrated using a vacuum manifold, resuspended in an equal volume of Milli-Q purified water, and then filtered through a 0.45-μm filter unit (polyvinylidene difluoride [PVDF] Titan syringe filter). Each sample was analyzed employing liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). The samples were compared to chemically synthesized c-di-GMP (Biolog) concentration standards ranging from 250 nM to 1.9 μM to determine the c-di-GMP concentration. Cyclic di-GMP was normalized to total soluble proteins. Soluble proteins were extracted by lysing a parallel cell sample with CelLytic B bacterial cell lysis and extraction reagent (Sigma). Cells were then mixed with an equal volume of <106-μm glass beads (Sigma) on a Vortex mixer with 10 repeated cycles of 1-min agitation with cooling intervals of 1 min. During the process, cells were maintained in a cold room or in ice water bath. The cell extract was collected after initial centrifugation at 1,000 × g for 1 min at 4°C and further centrifuged at 17,000 × g for 30 min at 4°C. The protein concentration was determined using a Pierce Bicinchoninic acid (BCA) protein assay kit.

The statistically significant effects of light conditions on c-di-GMP levels were determined using one-way analysis of variance (ANOVA) with a Fisher post hoc test using OpenStat statistical software (version 10.01.08; http://www.Statprograms4U.com [W. G. Milller]). Statistical analyses were performed utilizing 95% confidence intervals (P < 0.05).

**SUPPLEMENTAL MATERIAL**

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00451-13/-/DCSupplemental.

Figure S1, PDF file, 0.1 MB.
Table S1, DOCX file, 0.1 MB.
Table S2, DOCX file, 0.1 MB.
Table S3, DOCX file, 0.1 MB.

**ACKNOWLEDGMENTS**

This work was supported by the National Science Foundation (grant no. MCB-0643516 to B.L.M.), the National Institutes of Health (grant no. U54AI-057153 to C.M.W.), and support to Marco Agostoni from the U.S. Department of Energy (Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, grant no. DE-FG02-91ER20201 to B.L.M.).

We thank Shin-Han Shiu, Titus C. Brown, and Gaurav Moghe for useful discussions on bioinformatics analyses. Finally, we thank two anonymous reviewers for very helpful comments.

**REFERENCES**

1. Agrawal AA. 2001. Phenotypic plasticity in the interactions and evolution of species. Science 294:321–326.
2. Bailey S, Grossman A. 2008. Photoprotection in cyanobacteria: regulation of light harvesting. Photochem. Photobiol. 84:1410–1420.
3. Singh SP, Montgomery BL. 2011. Determining cell shape: adaptive regulation of cyanobacterial cellular differentiation and morphology. Trends Microbiol. 19:278–285.
4. Galperin MY, Higdon R, Kolker E. 2010. Interplay of heritage and habitat in the distribution of bacterial signal transduction systems. Mol. Biosyst. 6:721–728.
5. Tuckerman JR, Gonzalez G, Sousa EH, Wan X, Saito JA, Alam M, Gilles-Gonzalez MA. 2009. An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. Biochemistry 48: 9764–9774.
6. Galperin MY, Nikolskaya AN, Koonin EV. 2001. Novel domains of the prokaryotic two-component signal transduction systems. FEMS Microbiol. Lett. 203:11–21.
7. Schmidt AJ, Ryjenkov DA, Gomelsky M. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. J. Bacteriol. 187:4774–4781.
8. Dow JM, Fouchy Y, Lucey JF, Ryan RP. 2006. The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. Mol. Plant Microbe Interact. 19:1378–1384.
9. Thomas C, Andersson CR, Canales SR, Golden SS. 2004. PdB, a factor that stimulates psbA expression in the cyanobacterium *Synechococcus elongatus* PCC 7942. Microbiology 150:1031–1040.
10. Suzuki K, Babiitzke P, Kushner SR, Hobley L, Fung RK, Lambert C, Harris MA, Dabhi JM, King SS, Xanthomonas campestris. 2008. The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. Mol. Plant Microbe Interact. 19:1378–1384.
11. Thomas C, Andersson CR, Canales SR, Golden SS. 2004. PdB, a factor that stimulates psbA expression in the cyanobacterium *Synechococcus elongatus* PCC 7942. Microbiology 150:1031–1040.
12. Duerig A, Abel S, Bolcher M, Ncollier M, Schrvede T, Amiot N, Giese B, Jenal U. 2009. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. Genes Dev. 23:93–104.
13. He YW, Zhang LH. 2008. Quorum sensing and virulence regulation in *Xanthomonas campestris*. FEMS Microbiol. Rev. 32:842–857.
14. Hobley L, Fung RK, Lambert C, Harris MA, Dabhi JM, King SS, Basford SM, Uchida K, Till R, Ahmad R, Aizawa S, Gomelsky M, Sackett RE. 2012. Discrete cyclic di-GMP-dependent control of bacterial predation versus axenic Growth in *Bdellovibrio bacteriovorus*. PLoS Pathog. 8:e1002493.
15. Weber H, Pesavento C, Possling A, Tischendorf G, Hengeg R. 2006. Cyclic-di-GMP-mediated signalling within the sigma(n) network of *Escherichia coli*. Mol. Microbiol. 62:1014–1034.
16. Neumuebel MR, Golden JW. 2008. The *Anaeraba* sp. strain PCC 7120 gene all2874 encodes a diguanylate cyclase and is required for normal motility of *Synechocystis* sp. PCC 6803. Mol. Microbiol. 85:239–251.
17. Wierzchos J, Ascaso C, McKay CP. 2006. Endolithic cyanobacteria in halite rocks from the hyperarid core of the Atacama Desert. Astrobiology 6:415–422.

18. Jungblut AD, Lovejoy C, Vincent WF. 2010. Global distribution of cyanobacterial ecotypes in the cold biosphere. ISME J. 4:191–202.

19. Pan Z, Livoti E, Losi A, Gatzon MM, Ward DM. 2003. Geographical isolation in hot spring cyanobacteria. Environ. Microbiol. 5:650–659.

20. Cavalieri-Smith T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int. J. Syst. Evol. Microbiol. 52: 297–354.

21. Fulkowski PG, Knoll AH. 2007. Evolution of primary producers in the sea. Elsevier Academic Press, Burlington, MA.

22. Van der Horst MA, Hellingwerf KJ. 2004. Photoreceptor proteins, “star actors of modern times”: a review of the functional dynamics in the structure of representative members of six different photoreceptor families. J. Bacteriol. 186:37–130.

23. Montgomery BL. 2007. Sensing the light: photoreceptive systems and signal transduction in cyanobacteria. Mol. Microbiol. 64:16–27.

24. Romling U, Galperin MY, Gomelsky M. 2006. BLUF: a novel FAD-binding domain in archaea. Trends Biochem. Sci. 21:497–500.

25. Römling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol. Mol. Biol. Rev. 77:1–52.

26. Capone DG, Zehr JP, Parrell HW, Bergman B, Carpenter EJ. 1997. Trichodesmium, a globally significant marine cyanobacterium. Science 276:2121–2129.

27. Song JY, Cho HS, Cho JI, Jeon JS, Lagarias JC, Park YI. 2005. Global analysis of circadian expression in the cyanobacterium Synechocystis sp. strain PCC 6803. J. Bacteriol. 187:2465–2472.

28. Kucho K, Okamoto K, Tsuchiya Y, Nomura S, Nango M, Kanehisa M, Yashita H, Page L, Ramakrishna P, Satoh S, Sattely WM, Shimada Y, Taylor HL, Tomo T, Tsuchiya T, Wang ZT, Raymond J, Minuro M, Blankenship RE, Touchman JW. 2008. Niche adaptation and genome expansion in the chlorophyll d-producing cyanobacterium Acaryochloris marina. Proc. Natl. Acad. Sci. U. S. A. 105:2005–2010.

29. Gupta HK, Singh A, Sharma R. 2011. Genome sequence of Idiomarina loihiensis strain A28L, isolated from Pongang Lake, India. J. Bacteriol. 193:5875–5876.

30. Tandeau de Marsac N. 1977. Occurrence and nature of chromatic adaptation in Cyanobacteria. J. Bacteriol. 136:82–91.

31. Wilson A, Ajlani G, Verbaatym JM, Vass I, Kerfeld CA, Kirilovsky D. 2006. A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria. Plant Cell 18:992–1007.

32. Terauchi K, Ohmori M. 2004. Blue light stimulates cyanobacterial motility via a cAMP signal transduction system. Mol. Microbiol. 52:303–309.

33. Karpowicz SJ, Prochnik SE, Grossman AR, Merchant SS. 2011. The GreenC2t2 resource, a phylogenomically derived inventory of proteins specific to the plant lineage. J. Bacteriol. 193:21472–21499.

34. Meeks JC. 1998. Symbiosis between nitrogen-fixing cyanobacteria and plants. Bioscience 48:266–276.

35. Welsh EA, Liberton M, Stockel J, Loh T, Elvigtal T, Wang C, Wollam A, Fulton RS, Clifton SW, Jacobs JM, Aurora R, Ghosh BK, Sherman LA, Smith RD, Wilson RK, Pakrasi HB. 2008. The genome of Cyanotheris 51124, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. Proc. Natl. Acad. Sci. U. S. A. 105:15094–15099.

36. Loza V, Berrendero E, Perona E, Mateos P. 2013. Polyphasic characterisation of the benthic cyanobacterial diversity from biofilms of the Guadarrama river (Spain): Morphological, molecular, and ecological approaches. J. Phycol. 49:282–297.

37. Roeslers G, Huisjes EH, van Loosdrecht MC, Muyzer G. 2008. Diversity and expression of cyanobacterial hupS genes in pure cultures and in a nitrogen-limited phototrophic biofilm. FEMS Microbiol. Ecol. 63:292–300.

38. Massie JP, Reynolds EL, Koestler BJ, Cong JP, Agostoni M, Waters CM. 2012. Quantification of high-specificity cyclic diguanylate signaling. Proc. Natl. Acad. Sci. U. S. A. 109:12746–12751.

39. Davey ME, O’Toole GA. 2000. Microbial biofilms: from ecology to molecular genetics. Microbiol. Mol. Biol. Rev. 64:847–867.

40. Xu M, Su Z. 2009. Computational prediction of cAMP receptor protein (CRP) binding sites in cyanobacterial genomes. BMC Genomics 10:286.

41. Partensky F, Garzcezark L. 2010. Prochlorococcus: advantages and limits of pigmentation. Annu. Rev. Mar. Sci. 2:305–331.

42. Schirmmeister BE, Antonelli A, Bagheri HC. 2011. The origin of multicellularity in Cyanobacteria. BMC Evol. Biol. 11:45.

43. Honda D, Yokota A, Sugiyama J. 1999. Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine Synechococcus strains. J. Mol. Evol. 48:723–739.

44. Mobberley JM, Ortega MC, Foster JS. 2012. Comparative microbial metabolic analyses of modern marine thrombotic mats by barcoded pyrosequencing. Environ. Microbiol. 14:82–100.

45. Schatz D, Nagar E, Sendersky E, Parnasa R, Zilberman S, Carmeli S, Mastay Y, Shimon I, Klein E, Yeger O, Reich Z, Schwarz R. 2013. Self-suppression of biofilm formation in the cyanobacterium Synechococcus elongatus. Environ. Microbiol. 15:1786–1794.

46. Nakao M, Okamoto S, Kohara M, Fujihiro T, Fujisawa T, Sato S,
Tabata S, Kaneko T, Nakamura Y. 2010. CyanoBase: the cyanobacteria genome database update 2010. Nucleic Acids Res. 38:D379–D381.

KiriHina O, Fetherston JD, Bobrov AG, Abney J, Perry RD. 2004. HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in Yersinia pestis. Mol. Microbiol. 54:75–88.

Malone JG, Williams R, Christen M, Jenal U, Spiers AJ, Rainey PB. 2007. The structure–function relationship of WspR, a Pseudomonas fluorescens response regulator with a GGDEF output domain. Microbiology 153:980–994.

Malone JG, Abney J, Perry RD. 2004. HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in Yersinia pestis. Mol. Microbiol. 54:75–88.

Kirillina O, Fetherston JD, Bobrov AG, Abney J, Perry RD. 2004. HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in Yersinia pestis. Mol. Microbiol. 54:75–88.

65. Malone JG, Williams R, Christen M, Jenal U, Spiers AJ, Rainey PB. 2007. The structure–function relationship of WspR, a Pseudomonas fluorescens response regulator with a GGDEF output domain. Microbiology 153:980–994.

66. Malone JG, Williams R, Christen M, Jenal U, Spiers AJ, Rainey PB. 2007. The structure–function relationship of WspR, a Pseudomonas fluorescens response regulator with a GGDEF output domain. Microbiology 153:980–994.

67. Lovering AL, Capeness MJ, Lambert C, Hobley I, Sockett RE. 2011. The structure of an unconventional HD-GYP protein from Bdellovibrio reveals the roles of conserved residues in this class of cyclic-di-GMP phosphodiesterases. mBio 2:e00163-11.

68. Pathak GP, Losi A, Gärnter W. 2012. Metagenome-based screening reveals worldwide distribution of LOV-domain proteins. Photochem. Photobiol. 88:107–118.

69. Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, Wright F. 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics 25:126–127.

70. Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18:691–699.

71. Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160–174.

72. Cobley JG, Zerweck E, Reyes R, Mody A, Seludo-Unson JR, Jaeger H, Weerasuriya S, Navankasattus S. 1993. Construction of shuttle plasmids which can be efficiently mobilized from Escherichia coli into the chromatically adapting cyanobacterium, Fremyella diplosiphon. Plasmid 30:90–105.

73. Allen MM. 1968. Simple conditions for growth of unicellular blue-green algae on plates. J. Phycol. 4:1–4.