



HAL
open science

Light color regulation of photosynthetic antennae biogenesis in marine phytoplankton

David M. Kehoe, Avijit Biswas, Bo Chen, Louison Dufour, Théophile Grébert,
Allissa M. Haney, Kes Lynn Joseph, Indika Kumarapperuma, Adam A. Nguyen,
Morgane Ratin, et al.

► **To cite this version:**

David M. Kehoe, Avijit Biswas, Bo Chen, Louison Dufour, Théophile Grébert, et al.. Light color regulation of photosynthetic antennae biogenesis in marine phytoplankton. *Plant and Cell Physiology*, 2024, <10.1093/pcp/pcae115>. <hal-04725332>

HAL Id: hal-04725332

<https://hal.science/hal-04725332v1>

Submitted on 8 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Light Color Regulation of Photosynthetic Antennae Biogenesis in Marine Phytoplankton

Short title: Chromatic Acclimation in Marine *Synechococcus*

David M. Kehoe^{1*}, Avijit Biswas², Bo Chen¹, Louison Dufour³, Théophile Grébert³, Allissa M. Haney¹, Kes Lynn Joseph², Indika Kumarapperuma⁴, Adam A. Nguyen², Morgane Ratin³, Joseph E. Sanfilippo¹, Animesh Shukla¹, Laurence Garczarek³, Xiaojing Yang⁴, Wendy M. Schlachter² and Frédéric Partensky³

¹Department of Biology, Indiana University, Bloomington, Indiana, 47405, USA

²Department of Biological Sciences, Department of Chemistry, University of New Orleans, New Orleans, Louisiana, 70148, USA

³Sorbonne Université, Centre National de la Recherche Scientifique, UMR 7144 Adaptation and Diversity in the Marine Environment, Station Biologique, Roscoff 29680, France

⁴Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois, 60612, USA

*Corresponding Author
Email: dkehoe@iu.edu
FAX: +1 (812) 855-6705

ABSTRACT

Photosynthesis in the world's oceans is primarily conducted by phytoplankton, microorganisms that use many different pigments for light capture. *Synechococcus* is a unicellular cyanobacterium estimated to be the second most abundant marine phototroph, with a global population of 7×10^{26} cells. This group's success is partly due to the pigment diversity in their photosynthetic light harvesting antennae, which maximize photon capture for photosynthesis. Many *Synechococcus* isolates adjust their antennae composition in response to shifts in the blue:green ratio of ambient light. This response was named Type 4 chromatic acclimation (CA4). Research has made significant progress in understanding CA4 across scales, from its global ecological importance to its molecular mechanisms. Two forms of CA4 exist, each correlated with the occurrence of one of two distinct but related genomic islands. Several genes in these islands are differentially transcribed by the ambient blue:green light ratio. The encoded proteins control the addition of different pigments to the antennae proteins in blue versus green light, altering their absorption characteristics to maximize photon capture. These genes are regulated by several putative transcription factors also encoded in the genomic islands. Ecologically, CA4 is the most abundant of marine *Synechococcus* pigment types, occurring in over 40% of the population oceanwide. It predominates at higher latitudes and at depth, suggesting that CA4 is most beneficial under sub-saturating photosynthetic light irradiances. Future CA4 research will further clarify the ecological role of CA4 and the molecular mechanisms controlling this globally important form of phenotypic plasticity.

Key Words:

chromatic acclimation
cyanobacteria
light regulation
marine biology
photosynthesis
phycobilisome

INTRODUCTION

Our understanding of light color acclimation in picophytoplankton began with astute observations that were made by scientists in the early 1900s (Engelmann, 1902, Gaidukov, 1902, Gaidukov, 1903). These workers discovered that a subset of the filamentous cyanobacteria isolated from European freshwater lakes were capable of reversibly altering their color phenotype in response to the spectral composition of the ambient light in which they were grown. This color shift was not subtle: the cells were brick red in color when grown in predominantly green light and blue-green when the incident light color was primarily red (Fig 1A, left). Shifts between these two different color phenotypes was typically completed within five to seven days. Since these cells were red in green light and green in red light, this response was termed "Complementary Chromatic Adaptation" or CCA. However, studies over the past 120 years have demonstrated this name to be inaccurate and misleading for several reasons. First, "adaptation" requires a change in the genetic makeup of an organism, whereas "acclimation" involves only a change in the expression of an organism's genes. CCA has been demonstrated to involve only the latter (Sanfilippo et al., 2019a) and is actually a form of phenotypic plasticity, which are responses to changing environmental conditions that generally include rapid and reversible changes in an organism's morphology and physiology. Second, additional investigations focused on light color-driven changes in cyanobacterial coloration has shown that, in addition to the process described in the early 1900s, many cyanobacteria respond to a range of additional ambient light colors and display many different color phenotypes beyond the red and blue-green colors described above. Additionally, in these more recently discovered responses, cells will not change to the "complementary" color if provided an ambient light color matching their current color phenotype, making "complementary" an inaccurate term for all but the originally discovered red/green light response. Because of these additional findings, this process is now called "chromatic acclimation" (CA) and each of the different types given a numerical designation, based mainly on the order of their discovery (Sanfilippo et al., 2019a, Hirose et al., 2019a, Hirose et al., 2019b). The form of CA that is shown in Fig 1A and was originally termed "CCA" is now called "CA3".

The fact that different kinds of CA exist was first demonstrated by Tandeau de Marsac in 1977 (Tandeau de Marsac, 1977). By definition, all types of CA involve ambient light color changes that lead to alterations in the composition of the open chain tetrapyrrole (bilin) chromophores, or sometimes the proteins, that make up the photosynthetic light harvesting antennae found in most cyanobacteria. These structures, called phycobilisomes (PBS), are large, water soluble, macromolecular structures that are appressed to the cytosolic sides of the photosynthetic or thylakoid membranes within cyanobacterial cells. While recent studies have emphasized the variation that exists in PBS structures (Hirose et al., 2008, Kondo et al., 2005, Kondo et al., 2007, Deng et al., 2012), the majority of the known types of CA alter the composition of fan-shaped, or hemidiscoidal, PBS that are commonly found in the Cyanobacteria (Fig 1B) (Sanfilippo et al., 2019a). The changes that occur during CA help to maximize the frequency of photon capture by the PBS and thus the rates of photosynthesis (Campbell, 1996), which can be advantageous under conditions where the photosynthetic activity in a cell is not saturated.

This review will focus on the recent progress in understanding Type 4 CA (CA4), a form of CA that is only known to occur in marine *Synechococcus*, a polyphyletic group of unicellular picocyanobacteria. Periodically, the mechanisms and regulation of CA4 will be compared to and contrasted with that of CA3, the red-green form of CA described in the first paragraph and arguably the best characterized type of CA to date. The juxtaposition of these two CA systems provides a fascinating example of convergent evolution that has occurred in very different cyanobacterial genera to achieve a common goal: maximizing photon capture in sub-optimal irradiance situations.

DISCOVERY AND INITIAL CHARACTERIZATION OF CA4

The discovery and early investigations of the marine picophytoplankton *Synechococcus* spp. revealed that these unicellular cyanobacteria use bilins and hemidiscoidal PBS for photosynthetic light harvesting and, in addition to their sub-micrometer diameters, possessed quite small genomes, averaging 2.5 Mbp in size (Doré et al., 2020). These traits were in stark contrast to the 5 μm cell length and 10.2 Mbp genome size of the freshwater, filamentous cyanobacterium *Fremyella diplosiphon* (Yerrapragada et al., 2015), the model organism for the study of CA3. Even though marine *Synechococcus* possessed PBS, the small size and streamlined genomes of this group made it seem unlikely for them to be capable of CA. This assumption was shattered in 2001 by Brian Palenik's surprising discovery

of a completely novel form of CA that was responsive to blue and green light (Palenik, 2001). This response, which was designated CA4 five years later (Everroad et al., 2006), was the first type of CA found to respond to ambient light colors other than red and green. In the 1950s, action spectra were obtained showing that red and green light were the maximally effective wavelengths for shifting between the two possible states of CA3 (Hattori and Fujita, 1959, Fujita and Hattori, 1960, Fujita and Hattori, 1962, Diakoff and Scheibe, 1973, Haury and Bogorad, 1977, Vogelmann and Scheibe, 1978, Ohki et al., 1982). However, action spectra have not been generated for CA4, so although this response is clearly controlled by the blue and green light regions of the spectrum, we do not know yet the specific wavelengths that elicit the response in either direction. The possibility that either the blue or green light response is simply due to the absence of the other light color has also not been eliminated. This is an issue that becomes important when considering the possible photoreceptor(s) controlling CA4, as will be discussed in a subsequent section.

The basic structural features of marine *Synechococcus* PBS that are necessary to understand CA4 are described here, with more detailed information being provided in several excellent reviews (MacColl, 1998, Watanabe and Ikeuchi, 2013). All known PBS of this group are fan shaped or hemidiscoidal (Fig 1B), consisting of a central core and rods extending from the core. The rods and core are both built of cylinders containing two protein classes, phycobiliproteins and linker proteins. Photon energy is predominantly captured by phycobiliproteins in the rods and is transferred with high efficiency to photosynthetic reaction centers. These proteins bind one to three chromophore types that are covalently attached through cysteine residues and absorb specific, narrow bandwidths of light (Glazer, 1989). Linker proteins facilitate PBS assembly and attachment to the thylakoid membrane as well as tune the absorption of the phycobiliproteins and some possess chromophores (Tandeau de Marsac and Cohen-Bazire, 1977, Liu et al., 2005). Both rod and core phycobiliproteins are heterodimers consisting of α and β subunits. The specific cysteine residues that are the sites of bilin attachment in the α and β subunits are chromophorylated by phycobilin lyases, a class of enzymes required for the covalent bilin attachment (Schluchter et al., 2010). The stability of the α/β heterodimers depends on the attachment of these chromophores (Scheer and Zhao, 2008, Wiethaus et al., 2010, Anderson and Toole, 1998). Three fully chromophorylated heterodimers associate to form a ring, and two of these rings are stacked to form another, thicker ring of the same diameter that is called a disc. From three to six heterodimeric hexamers are further stacked to form a rod, with a series of specific linker proteins inserted into the centers of two adjacent discs to maintain their association (Six et al., 2005, Six et al., 2007). Overall, the phycobiliproteins and linker proteins within the PBS can be thought of as providing the necessary three-dimensional structure and local environments for the chromophores to absorb and correctly transfer their absorbed photon energy to the thylakoid membrane (Glazer, 1989, Bryant and Gisriel, 2024).

The rods of marine *Synechococcus* PBS may contain one, two, or three different types of phycobiliproteins, depending on the strain (Fig 1B), and these different rod compositions led Six and co-workers to classify the corresponding pigment phenotypes (PT) as PT1, PT2, and PT3, respectively (Six et al., 2007). The rod discs that are closest to the core always contain phycocyanin (PC), which possesses the bilin chromophore phycocyanobilin (PCB) and has an absorbance maximum (λ_{\max}) in the red region, at approximately 620 nm. In *Synechococcus* PT1 isolates, the more distal discs also contain PC, resulting in PBS that are well-suited for absorbing red light. However, the narrow absorbance range of PC means that isolates with such PBS poorly absorb ambient green or blue light. PT2 isolates contain a phycobiliprotein historically called phycoerythrin I (PE I) in the two to four discs that are beyond the core-proximal PC disc (Six et al., 2007). Like other phycobiliproteins, PE-I is an α/β heterodimer, and these two subunits are encoded by the *cpeBA* operon. These gene designations have led to the PE-I α and PE-I β subunits commonly being called CpeA and CpeB. In PT2 isolates, these subunits covalently attach the green-absorbing phycoerythrobilin (PEB) (λ_{\max} 550 nm), with two PEB attached to the α subunit and three attached to the β subunit (Six et al., 2007, Humily et al., 2013). In addition to possessing PC and PE-I, PT3 isolates have PE-II (Six et al., 2007, Ong et al., 1984, Ong and Glazer, 1991). The sequences of the α and β subunits of this phycobiliprotein are very similar to those of PE-I, and discs made of PE-II are always located at the core-distal ends of the rods. The genes encoding the PE-II subunits are also in an operon and called *mpeA* and *mpeB*. The proteins they encode, MpeA and MpeB, each have three covalently attached bilins, which can be either PEB or the blue-light absorbing phycourobilin (PUB) (λ_{\max} 495 nm) (Six et al., 2007, Humily et al., 2013). PBS that contain both PE-I and PE-II have a greater number of discs per rod than those with PE-I only and thus greater potential for photon capture.

The initial identification of marine *Synechococcus* capable of CA4 was carried out by growing isolates in either blue or green light for several weeks, then identifying those with pigmentation differences in these two light colors using fluorescence excitation, which is superior to absorbance measurements due to its ability to detect PEB and PUB with great specificity and sensitivity (Six et al., 2007). By measuring the ratio of the fluorescence excitation maxima of PUB and PEB at 495 nm and 545 nm, respectively ($Exc_{495:545}$), four subtypes were initially delineated within PT3 (Six et al., 2007). Those with an $Exc_{495:545}$ of approximately 0.4 were designated subtype 3a. This ratio was unchanged when these cells were grown in either green or blue light, and because of their high PEB content, PT 3a cells are called "green light specialists" (Sanfilippo et al., 2016). Isolates with an $Exc_{495:545}$ values between 0.6 and 1.6 that were also unchanged during growth in green or blue light were designated subtype 3b but are not further considered here because they have been found to be relatively rare. Those with an $Exc_{495:545}$ of approximately 1.6 were designated subtype 3c and, as for subtype 3a, this ratio was not altered when cells were grown in blue or green light. Because these isolates consistently contain relatively high amounts of PUB, they have been referred to as "blue light specialists" (Sanfilippo et al., 2016). The fourth subtype, 3d, consisted of isolates that displayed an $Exc_{495:545}$ of approximately 0.6 when grown in green light and an $Exc_{495:545}$ of approximately 1.6 when grown in blue light (Fig 1C). Thus, 3d cells somewhat phenocopied the 3c subtype when grown in blue light and the 3a subtype when grown in green light, although they could not reach as low of an $Exc_{495:545}$ value as the 3a subtype in green light. These phenotypes were "plastic", that is, the isolates could change between phenotypes within three to five days after being switched from one ambient light color to the other (Everroad et al., 2006, Humily et al., 2013, Dufour et al., 2024). The isolates were designated "Type 4 Chromatic Acclimators" (CA4 strains) or "blue-green generalists" (Everroad et al., 2006, Sanfilippo et al., 2016, Shukla et al., 2012), and are phenotypically pink when grown in green light and phenotypically orange when grown in blue light (Fig 1A, right).

INITIAL UNDERSTANDING OF CA4 MOLECULAR PROCESSES

The first several lines of evidence that the molecular basis of CA4 was unique and different from other known types of CA, including CA3, was provided by the analysis of the phycobiliprotein composition within the PBS from marine *Synechococcus* grown in either blue or green light. These studies found that there was no detectable difference in PBS protein composition between these two conditions (Everroad et al., 2006, Shukla et al., 2012), which was quite different from previously described forms of CA, which involve changes in the makeup of both bilins and phycobiliproteins within the PBS under different light color conditions (Sanfilippo et al., 2019a). These studies also demonstrated that in two different marine *Synechococcus* strains, the changes in bilin attachment during CA4 were unique and occurred at three of the 11 cysteines used for chromophore attachment in PE-I and PE-II (Shukla et al., 2012, Grébert et al., 2021). These locations are cysteine 139 of CpeA (CpeA-C139), MpeA-C83, and MpeA-C140. Each of these three sites has PEB attached during growth in green light and PUB attached during growth in blue light (Everroad et al., 2006, Shukla et al., 2012). Because every PBS rod has multiple copies of each of these proteins, this means that there are 288 PUB-PEB chromophore changes per PBS that occur during CA4. In other forms of CA, the only shifts in bilins that were known to occur were between PCB and PEB (Sanfilippo et al., 2019a).

Efforts to understand the cellular mechanisms through which CA4 operates began by identification of the genes encoding the components required for this response. This was accomplished using two approaches. Initially, the genomes of many of the marine *Synechococcus* isolates that were selected to establish the PT 3 subtypes were sequenced, and comparative genomics approaches were used to identify any genomic regions that were unique to 3d isolates. This method was very productive, leading to the identification of a DNA region approximately 5 kbp in size with characteristics of a genomic island (GI) that was found only in CA4 strains (Six et al., 2007). Subsequent analyses of a larger sized group of marine *Synechococcus* isolates expanded this finding with the discovery that two distinct forms of the GI existed in the CA4 population (Humily et al., 2013). The first discovered was named "CA4-A" and the second was named "CA4-B", and the corresponding PT 3 subtypes were called 3dA and 3dB, respectively. Interestingly, the pigmentation phenotypes in monochromatic blue and green light for isolates containing these two different GIs were found to be essentially the same. However, a recent study showed that 3dA and 3dB strains display significantly different $Exc_{495:545}$ values when grown in a mixture of blue and green light, as well as different acclimation kinetics after shifts between blue and green light conditions (Dufour et al., 2024).

While the gene content within each type of GI varied slightly, there are five genes that were consistently found within the CA4-A GI and four genes that were invariably present in CA4-B GIs. The two GIs, which contain some overlap in the genes they possess, are diagrammed in Fig 2. Some of the predicted proteins encoded within these GIs provided important clues about the mechanism and regulation of CA4, while others were unlike any proteins characterized to date. Overall, however, the discovery of these GIs enabled subsequent investigations to be tightly focused on the roles of their encoded gene products on CA4. The conclusion that these GIs were either primarily or solely responsible for the CA4 response was further supported by RNAseq data obtained from the CA4-A model organism *Synechococcus* sp. RS9916 (hereafter 9916), which showed that RNAs from only three genes showed significantly different levels between cells grown in blue and green light, and all were within the CA4-A GI: *mpeZ* (35x increase in blue light), *fcic* (10x increase in blue light), and *unk10* (now called *bbaA*) (40x increase in blue light) (Sanfilippo et al., 2016). Individual fusions of the *fcic* and *mpeZ* promoters to the *gusA* reporter gene demonstrated that the CA4-mediated changes in RNA levels for these two genes was at the level of transcription (Haney and Kehoe, unpublished results).

The proteins encoded within these two GIs displayed intriguing similarities as well as differences. In the CA4-A GI, *mpeZ* was predicted to encode a protein with similarity to phycobilin lyases and to the protein encoded by *mpeW* in the CA4-B GI. Both GIs contained the *fcia* and *fcib* genes, which were similar in their C-terminal regions to the DNA binding motifs of AraC family transcription factors. The small open reading frame *fcic*, which was present in all CA4-A GIs but not in any CA4-B GIs, encoded a protein related to the ribbon-helix-helix domain of some phage repressors (Aravind et al., 2005). All CA4-A and CA4-B GIs contain *unk10*, which encodes a small protein of unknown function. The analysis of Unk10 function has led to its being renamed BbaA, as explained later in this review.

GENETIC AND BIOCHEMICAL ANALYSES OF THE CA4 MECHANISM

As described above, the use of comparative genomics and RNAseq analyses provided important insights into the molecular basis of CA4. It suggested that this response could be acquired by the acquisition of one of two CA4 GIs, that it involved differential expression of at least one bilin lyase family member, and that it was controlled by several transcription factors. Yet no significant additional advances were possible using only these approaches. Further advances were made through the development of methods that allowed culturing, plating, and genetically modifying CA4-capable marine *Synechococcus*. Initially, disruption and overexpression of CA4-A GI genes were achieved in 9916 using the suicide and autonomously replicating plasmids, respectively, and methods for transforming non-CA4 marine *Synechococcus* cells developed by Brahmasha (Brahmasha, 1996). Subsequent development of additional molecular genetic tools for marine *Synechococcus*, including CRISPR-based mutagenesis, reporter gene systems and overexpression systems, in conjunction with molecular biological and biochemical approaches, has begun to allow the identification of the functions of the proteins encoded by the genes in both CA4-A and CA4-B GIs.

The first CA4-A GI gene to be disrupted in 9916 was *mpeZ*, since its sequence strongly suggested that it was a bilin lyase. Surprisingly, mass spectrometry analysis of PBS isolated from a *mpeZ* disruption mutant showed that, compared to control cells, only MpeA-C83 chromophorylation was affected, with this residue binding PEB instead of PUB during growth in blue light (Shukla et al., 2012). In addition, expression of MpeZ in *Escherichia coli* cells, along with the genes needed to produce PEB and MpeA, demonstrated that MpeZ was a member of a subset of the bilin lyase family called lyase-isomerases. Specifically, MpeZ binds PEB and isomerizes it to PUB during the attachment to MpeA-C83. This explained only one of the three chromophore changes occurring during CA4. It also raised the question of what enzyme was adding PEB to MpeA-C83 of the *mpeZ* mutant in blue light, and presumably green light as well. Survey of the 9916 genome coding sequences identified *mpeY* as the gene encoding the protein with the greatest sequence similarity to MpeZ. An *mpeY* disruption mutant was generated and analyzed for its PBS pigment content during CA4. While wild type chromophorylation patterns were found at two of the three residues affected by CA4, MpeA-C83 was found to have PUB covalently attached in both blue and green light (Sanfilippo et al., 2019b). Overexpression of *mpeY* led to PEB attachment at this residue under both light conditions. Analysis of RNA-seq data demonstrated that while *mpeY* RNA levels were not regulated by CA4, they were higher than *mpeZ* RNA levels in 9916 during growth in green light. In addition, the increased abundance of *mpeZ* transcripts in blue light resulted in *mpeZ* RNA levels that were significantly elevated over those from *mpeY*. Therefore, it is very likely that the

chromophore that is attached to MpeA-C83 is determined by the relative levels of MpeY and MpeZ under the two light color conditions (Sanfilippo et al., 2019b).

The CA4-mediated chromophorylation of MpeA-C140 and CpeA-C139 have been a more difficult issue to address, since no other protein with similarity to known lyase family members is encoded within the 9916 CA4-A GI (Fig 2). The similarity of FciA, FciB, and FciC to known DNA binding proteins made it unlikely that these three proteins were acting as lyases, leaving Unk10/BbaA as the only likely candidate for this role, despite its small (12 kDa) size.

The sequence of Unk10/BbaA has strong sequence similarity to Nif11 domains (Pfam078620) (Jacobson et al., 1989, Haft et al., 2010), which have been linked to nitrogen fixation and secondary metabolite production and are found primarily in cyanobacteria and proteobacteria. The *unk10/bbaA* gene is present in all blue light specialists and appears to be very ancient, as it is found in phylogenetically distant marine *Synechococcus* lineages (Grébert et al., 2022). Initial results from the CRISPR-mediated deletion of this gene indicate that Unk10 is involved in adding PUB to phycobiliproteins during growth in blue light (Chen et al., unpublished results). While this finding is being further examined, the name of this gene has been provisionally changed from *unk10* (for "unknown" protein) to *bbaA* ("blue light bilin addition") to reflect these results. BbaA's mechanism of operation is not yet known.

Thus, the actual mechanism underlying the CA4-A, and likely the CA4-B, response is highly streamlined, with increased production of only two different proteins being directly or indirectly responsible for the changes in chromophore composition between blue and green light. This mechanism is very simplified compared to the processes involved in the CA3 response in *F. diplosiphon*, which requires the differential production of at least 16 phycobiliproteins, linkers, and enzymes that are located in five separate operons throughout the genome (Sanfilippo et al., 2019a).

Studies of CA4-B have primarily used *Synechococcus* sp. A15-62 (hereafter called A15-62), which contains *mpeW* in addition to *bbaA* in its GI (Fig 2). The expression of *mpeW* was the opposite of *mpeZ* in 9916, with RNA levels being over 30x higher in green light than blue light (Humily et al., 2013). Interestingly, *bbaA* RNA levels were 40x lower in green light than in blue light. These results suggested that the enzymatic activities of MpeZ in 9916 and MpeW in A15-62 might be different. Deletion of *mpeW* in A15-62 and biochemical characterization of MpeW activity confirmed this, demonstrating that this enzyme is a bilin lyase, adding PEB to MpeA-C83 in green light (Grébert et al., 2021). Interestingly, the addition of PUB in blue light was found to be carried out by the lyase-isomerase MpeQ in A15-62. Thus, during CA4-B the MpeW/MpeQ enzyme pair operates conversely, but in much the same way, that the CA4-A MpeZ/MpeY enzyme pair functions in 9916. Although the role of BbaA has not yet been determined in A15-62, the similar differential expression of *bbaA* RNA in A15-62 and 9916 in blue and green light suggests that it has the same function in CA4-A and CA4-B strains.

These findings raise the question of why there are two types of CA4 GIs. It has been proposed that the CA4-A GI can convert a green light specialist, but not a blue light specialist, to a CA4 strain, while the CA4-B GI can convert a blue light specialist but not a green light specialist into a CA4 strain (Sanfilippo et al., 2019a). Our initial results indicate that transfer of a CA4-B GI into the two specialist types converts a blue light specialist, but not a green light specialist, into a CA4 strain (Grébert et al., unpublished results). This provides a reasonable explanation for why two forms of the CA4 GI exist and provides strong evidence that all the components required to confer CA4-B are encoded within the CA4-B GI.

THE REGULATION OF CA4

The ability of cyanobacteria to sense changes in ambient light color is one of the most fascinating aspects of CA. For those systems that have been defined, light color sensing has been shown to be mediated by members of the phytochrome superfamily (Sanfilippo et al., 2019a). These photoreceptors possess a covalently attached bilin chromophore which undergoes an isomerization in response to the absorption of a photon. The absorbance maxima for the two isomerization states of the bilin are different, allowing the bilin to be used efficiently as a sensor of the ratio of two wavelengths of light. This ratio is converted into a cellular signal by changing the activity of the attached

protein. Such output signals vary, but thus far for CA in cyanobacteria, they are changes in the kinase activity of the protein.

However, the genomes of marine *Synechococcus* such as 9916 and A15-62 do not contain any genes that encode proteins with any strong sequence similarity to members of the phytochrome superfamily. In addition, both CA4-A and CA4-B GIs contain genes encoding FciA and FciB, whose N-terminal domains do not show strong sequence similarity to any known proteins and whose C-terminal domains are related to DNA binding motifs of the AraC/XylS transcription factor family. The fact that the CA4-B GI is sufficient to confer CA4 to a blue light specialist, as described in the previous section, strongly supports the proposal that the sensory components for CA4 are encoded in the CA4 GIs and suggests that FciA and FciB have light color sensing roles.

Genetic approaches have provided a significant amount of information about the roles of FciA and FciB by interrupting either *fciA* or *fciB* in 9916 (Sanfilippo et al., 2016). These mutants had opposite phenotypes: the *fciA* mutant constitutively displayed the green light phenotype while the *fciB* mutant constitutively showed the blue light phenotype. In addition, overexpression of FciA made the cells hypersensitive to blue light while overexpressing FciB made them hypersensitive to green light. Taken together, these results demonstrate that FciA behaves as an activator while FciB behaves as a repressor. The mutant phenotypes were the result of constitutively low-level expression of the *fciC*, *mpeZ*, and *bbaA* genes in the *fciA* mutant and constitutively high-level expression of these genes in the *fciB* mutant. RNAseq data showed that *fciA* and *fciB* transcript levels did not change during growth in blue versus green light, suggesting that the regulation of the corresponding proteins' activity was occurring at the post-translational level. Although other explanations are possible, these results indicate that in CA4-A strains, FciA is acting as an activator in blue light while FciB is acting as a repressor in green light. No data are yet available to support the proposal that these two proteins are functioning as photoreceptors or whether they are acting as transcription factors, as their C-terminal domains suggest. The functions of FciA and FciB in CA4-B strains are being investigated in A15-62 but thus far, only *fciB* has been deleted and its characterization is still in progress (Dufour, Ratin et al., unpublished results). It is likely that the roles of both FciA and FciB are different in CA4-A and CA4-B strains.

The gene encoding FciC is only present in CA4-A GIs. Its 10x higher expression in blue light versus green light and loss of CA4 regulation in 9916 *fciA* and *fciB* mutants shows that it is under control of FciA and FciB. After its expression is increased upon a shift to blue light, FciC might or might not subsequently control the expression of the other CA4-A responsive genes *mpeZ* and *bbaA*. If it does control their expression, it could be solely responsible for their increase or could work with FciA and/or FciB to accomplish this. A CRISPR-mediated *fciC* deletion mutant is currently being analyzed and does appear to have a positive regulatory role in 9916. It is not clear why FciC should be needed in CA4-A strains but not in CA4-B strains. Perhaps the necessary level of gene expression in CA4-A strains can only be achieved through amplification of the signal through FciC.

Overall, models are beginning to become clear for the regulation of both CA4-A and CA4-B systems, even though the mechanistic details have not yet been worked out (Fig 2). For CA4-A, FciA appears to act as an activator of *fciC* expression in blue light, while FciB acts as a repressor in green light. We propose that a 10x increase in *fciC* RNA in blue light results in higher FciC levels, increasing the expression of *mpeZ* and *bbaA*. The proteins encoded by these two genes increase in blue light, leading to the attachment of PUB at the three residues that are regulated by CA4. Although the CA4 mechanisms still needs to be elucidated in CA4-B strains, we speculate that FciB could directly repress *mpeW* expression and increases *bbaA* expression in blue light, while FciA could act as an activator of *mpeW* expression and a repressor of *bbaA* expression during growth in green light. These straightforward regulatory schemes are much less complex than those controlling CA3 in *F. diplosiphon*, which involves two phytochrome superfamily members and their downstream signal transduction components, as well as the hierarchical regulation of two downstream operons by an activator that is encoded within another CA3-regulated operon (Sanfilippo et al., 2019a). Both the regulation and mechanisms controlling the CA4 and CA3 responses have clearly evolved independently, with many fewer components in the CA4 system. Yet their final responses are similar, altering the bilin chromophores within the PBS to optimize light capture for photosynthesis.

THE STRUCTURE, FUNCTION AND BIOCHEMISTRY OF LYASE ISOMERASES

An important part of the process of CA4 is the interplay between bilin lyases and bilin lyase-isomerases during growth in blue or green light. Therefore, substantial effort has been put into determining the structural basis that distinguishes between the lyase and lyase-isomerase functions of the MpeQ/MpeW and MpeZ/MpeY family. The lyases and lyase-isomerases in the MpeQWYZ family encoded in the genomes of all the PT 3 subtypes have 60-80% sequence identity, suggesting that evolutionary changes in specific residues lead to their divergent activities. To better understand the mechanisms separating lyase and lyase-isomerase functions, the Yang lab determined the crystal structure of the MpeQ lyase-isomerase at 2.5 Å resolution (Kumarapperuma et al., 2022). The single-chain structure of MpeQ revealed a large active-site chamber in a “question-mark”-shaped protein architecture (Fig 3A). Structural and sequence analysis suggest that this architecture is representative for bilin lyases in the E/F family of lyases and calls for the revision of the previously reported CpcE/CpcF model (Zhao et al., 2017). Based on this MpeQ structure and site-directed mutagenesis, several key residues responsible for the lyase and isomerase activities of MpeQ were identified. The structural analyses revealed a nucleophilic tyrosine (Y318 in MpeQ) within a catalytic triad in several bilin lyases, suggesting a shared reaction mechanism among bilin lyases of distinct protein scaffolds (Kumarapperuma et al., 2022) (Fig 3A).

The high sequence identity among the 106 sequences of this MpeWQYZ family was used to guide site-directed mutagenesis and identify the “swap” residues that convert one activity into another, using the established assays the Schlachter lab developed for recombinant MpeQ and MpeW (Grébert et al., 2021). The conserved residues that are unique to each enzyme group were the initial focus. For example, V319-T320 of MpeQ (and MpeZ) are conserved among lyase-isomerases whereas Gly/Ala occupy the same positions in lyases (MpeY and MpeW). V319 is located immediately next to the catalytic Y318 in MpeQ. The possible steric effects on isomerization by position 319 were explored by changing V319 to Gly, Ala, or Phe, producing a strong effect on the isomerization capability of MpeQ (Fig 3B). When bulkier side chains were present, isomerization took place, but when these residues were smaller (for example, Gly), more PEB was observed at C-83-MpeA. When G319 in MpeW was changed to bulkier residues (in conjunction with some swap sites), more isomerase activity was observed (Kumarapperuma et al., 2022).

However, the Val/Gly substitution alone was not sufficient to confer exclusive PEB attachment by MpeQ, even in combination with several swap sites. This suggests that a switch from a lyase to lyase-isomerase requires contributions from other swap sites, which together collectively create an active site that is spacious enough to allow the resolution of the steric conflict on the A-ring of PEB during the ligation reaction. In MpeW, there is enough space in the active site to allow the non-planar A-ring conformation of PEB, allowing its direct ligation. In MpeQ, the active site is less spacious, and PEB to PUB isomerization must occur to resolve the steric conflict with the A-ring of the bilin prior to ligation, resulting in PUB addition (Fig 3B). These data also suggested that the isomerase activity is an “add-on” function associated with the lyase reaction catalyzed by lyases which arises from collective steric effects from the swap residues near the active site. Together these results are beginning to clarify how these marine *Synechococcus* strains have evolved the capability to adapt to blue light and carry out CA4 through the addition of this “isomerase” activity.

THE EVOLUTION OF CA4

The absence of *fcIA* and *fcIB* homologs in all cyanobacteria other than marine *Synechococcus*, including *Crocospaera* spp., which also contain two PE types (encoded by *cpeBA*- and *mpeBA*-like genes) that bind both PEB and PUB, indicates that the CA4 process is highly specific to this group. This raises the question of where this unique trait came from. Surprisingly, both types of CA4-capable cells (3dA and 3dB) are found in two deep branches of “cluster 5”, the cyanobacterial group that encompasses all marine *Synechococcus* (Herdman et al., 2001) this radiation, namely in sub-clusters 5.1 and 5.3 (Dufresne et al., 2008). In contrast, most characterized members of sub-cluster 5.2 (Cyanobium; Doré et al. 2020) lack PE-II and all are incapable of CA4 (Humily et al., 2013, Xia et al., 2020, Grébert et al., 2022). This suggests that not only the advent of CA4, but also the differentiation between the two CA4 types, occurred before the divergence of sub-clusters 5.1 and 5.3, which is thought to have taken place about 400-880 million years ago, i.e. during the Neoproterozoic eon (Sánchez-Baracaldo et al., 2019).

Comparisons of the numerous available marine *Synechococcus* genomes, which cover the entire pigment diversity existing within this group, including many PT 3dA and 3dB representatives, showed that most genes involved in the synthesis of PBS rods are gathered into a large GI (Six et al., 2007, Grébert et al., 2022). This so-called "PBS rod region", whose gene content and organization are specific for the corresponding PT, always starts with a tRNA-Phe_{GAA}, which appears to be a recombination hotspot. In several *Synechococcus* genomes this tRNA specifically serves as an insertion site for exogenous mobile genetic elements such as recombinases and transposases. The hypervariable gene region that transiently hosts these mobile elements has hallmarks of tycheposons, a novel family of DNA transposons recently identified upstream some GIs in many *Prochlorococcus* genomes (Hackl et al., 2023). These tycheposons are thought to have a key role in niche adaptation. Besides carrying recombination and transposition genes that are otherwise scarce in picocyanobacteria genomes, they often carry functional genes as cargo that can remain in the genome after the disappearance of the mobile elements. In *Prochlorococcus*, this process was found to enable natural populations to cope with local nutrient deficiency by acquiring genes involved in the uptake and assimilation of nitrogen, phosphorus, or iron. The presence of tycheposons in *Synechococcus* genomes upstream the PBS rod region indicates that they played a crucial role in the adaptation of this organism to the five different spectral niches occurring in the ocean (Stomp et al., 2007, Holtrop et al., 2021), and notably in the acquisition of the ability to carry out CA4 (Grébert et al., 2022).

Interestingly, although the CA4-A type GI found in PT 3dA can be located anywhere in the genome, it is sometimes found within the 5'-end of the PBS region (Grébert et al., 2022), as observed for cargo genes sometime after their import via tycheposons (Hackl et al., 2023). In contrast, the CA4-B GI is always situated in the middle of the PBS rod region (Humily et al., 2013, Grébert et al., 2022). Intriguingly, the PBS genes of PT 3dA cells are phylogenetically distant from their homologs in PT 3a cells (green light specialists), and this divergent genetic background, as well as the different gene content of the PBS rod region, likely decreases the chances that a PT 3a cell could successfully acquire CA4 ability by insertion of a CA4-A GI. An example of this is *Synechococcus* sp. MVIR-18-1, which has a typical PT 3a genetic background and PBS rod region, also possesses a complete CA4-A GI yet is incapable of CA4 (Humily et al., 2013). In contrast, PBS genes from PT 3c (blue light specialists) and 3dB cells are so closely related phylogenetically that they cannot be discriminated from each other in phylogenetic trees (Grébert et al., 2018, Grébert et al., 2021). This suggests that if a CA4-B GI is imported into a PT 3c genome, for instance via a tycheposon mechanism, this lateral transfer has a relatively great chance of conferring CA4 ability to the PT 3c cell. The absence of CA4-B GIs in locations other than the middle of the PBS region also suggests that, once acquired, the CA4-B GI must be rapidly integrated at this specific location, possibly using the *unk10/bbaA* gene as a recombination site, since this is the only gene that is shared between the CA4-B GI and the PBS region of PT 3c cells.

THE ECOLOGY OF CA4

The discovery of CA4 was important because it occurred in marine *Synechococcus*, which is the second-most abundant phototroph on Earth, with an annual global mean abundance of 7×10^{26} cells (Flombaum et al., 2013). This raised the possibility that the process of CA4 could be of major importance to explain this ecological success. However, for quite some time after its discovery the prevalence of CA4 in natural *Synechococcus* populations globally was unclear. In recent years, some of the most important findings of CA4 research have provided insights into this issue by showing that CA4 is globally widespread in marine *Synechococcus* populations. This was a challenging issue to address because fluorescence methods such as flow cytometry, which is routinely used to enumerate picocyanobacteria in marine samples, cannot discriminate between blue light specialists and blue light-acclimated CA4-capable cells, and cannot reliably differentiate between green light specialists and green-acclimated CA4-capable cells, a problem that is even more difficult for samples in which several *Synechococcus* PTs co-occur (Olson et al., 1990). Humily *et al.* (Humily et al., 2014) were the first to assess the relative abundances of CA4-capable cells within natural *Synechococcus* populations by amplifying the *cpeBA* operon from a water sample collected off Roscoff, France. They found that 28% of all *cpeBA* amplicons could be assigned to PT 3dA and the rest to PT 3a (green light specialists). Using the same metabarcoding approach, Xia *et al.* (Xia et al., 2017) found that at several stations in the northwestern Pacific Ocean located above 40°N, all *cpeBA* reads could be assigned to PT 3dA, while this proportion tremendously decreased further south. Yet, this single-marker approach did not allow these authors to discriminate all PTs, particularly not PT 3c from 3dB since, as mentioned previously, their *cpeBA* sequences are too closely related to be discriminated. Grébert *et al.* (Grébert et al., 2018) overcame this problem by using a combination of three PBS

gene markers, *cpcBA*, *mpeBA* and *mpeW*, which proved to be effective in identifying all the major *Synechococcus* PTs. This strategy allowed them to precisely assess the proportion of 3dB cells which possess the *mpeW* gene within the total number of *mpeBA* sequences attributed to the combined number of 3c and 3dB cells. By retrieving reads of these three gene markers from metagenomes collected along the *Tara* Oceans transect, these authors constructed the first global map of *Synechococcus* PTs in the world's oceans (Fig 4B).

These analyses revealed the global significance of CA4 because CA4-capable cells accounted for 37.3% of all *Synechococcus* reads at the surface and 53.8% of all the reads at depth along the *Tara* Oceans transect (Fig 4A). Interestingly, the relative proportion of PT 3dA and 3dB was well balanced, as they globally represented 22.6% and 18.9 % of all *Synechococcus* reads, respectively. However, their distribution was very different, with PT 3dA cells mainly found in cold, nutrient-rich green waters at high latitudes and PT 3dB cells more abundant in nutrient-poor waters at low latitudes (Fig 4B and C) (Grébert et al., 2018). Another fascinating finding from this study was that *Synechococcus* populations thriving in warm, iron-depleted regions such as the South Pacific Ocean often displayed a PT 3dA genotype. Yet many of these cells were found to contain deletions in three genes found in wild type PT 3dA cells: *fciA*, *fciB* and *mpeY*. These naturally occurring mutations were also found in strain BIOS-E4-1, which displayed a constitutive blue light phenotype (Humily et al., 2013). In this strain the lyase-isomerase-encoding gene *mpeZ* and *bbaA/unk10* are still present but are no longer regulated by FciA and FciB, resulting in their constitutive expression and allowing MpeZ to add PUB at MpeA-C83 in all light color conditions. Thus, in low-iron areas, natural selection has seemingly favored the transformation of 3dA type CA4-capable cells into blue light specialists, which themselves derived a long time ago from green light specialists (see above). Ongoing research examining the ecology of CA4-capable *Synechococcus* is addressing their seasonal variation and depth distribution (Dufour, Faure et al., unpublished). Together with the results described above, these findings will help to model the global oceanic distribution of the three main PTs (green light specialists, blue light specialists, and CA4-capable cells) based on their differing absorption properties using the Darwin model (Dutkiewicz et al., 2015, Mattei et al., unpublished results).

SUMMARY AND CONCLUDING REMARKS

Our understanding of CA4 in marine *Synechococcus* has progressed remarkably since its discovery in 2001 (Palenik, 2001). Although many forms of CA have been described in cyanobacteria (Sanfilippo et al., 2019a, Hirose et al., 2019b, Hirose et al., 2019a), each with its own unique features and attributes, CA4 is different from other types of CA in several ways.

First, it is arguably the most widespread form of CA since it is estimated to occur in more than 40% of the 7×10^{26} marine *Synechococcus* cells worldwide (Flombaum et al., 2013, Grébert et al., 2018). A process as globally ubiquitous as CA4 must be considered when calculating photosynthetic efficiencies of picophytoplankton in the world's oceans. And although CA4 is considered a CA response, it may also be appropriate to think of it as a low light acclimation response, increasing the efficiency of photon capture under light-limiting conditions, explaining its frequency at higher latitudes and greater depths (see The Ecology of CA4 above).

Second, there are two distinctly different molecular processes and gene sets that are capable of conferring CA4, CA4-A and CA4-B, and these appear to be equally widespread in the natural environment (Six et al., 2007, Humily et al., 2013, Grébert et al., 2018). While the reasons for the co-existence of these two forms remain to be fully understood, it is attractive to posit that CA4-A converts green light specialists into CA4 strains while CA4-B converts blue light specialists into CA4 strains. If true, we are observing an intriguing snapshot of the evolution of a widespread and important form of phenotypic plasticity in the world's oceans. This raises the question of the dynamics of CA4 acquisition in marine *Synechococcus* cells. Are CA4-A and/or CA4-B strains becoming a greater percentage of the total *Synechococcus* population, or are their numbers static or even decreasing?

Third, the CA4 mechanism that solely relies on the interplay between lyases and lyase-isomerases to control PBS chromophore composition in blue versus green light represents a unique approach to carrying out the process of CA. In other forms of CA, light color-regulated expression of genes encoding PBS proteins and bilin biosynthetic enzymes are the norm (Sanfilippo et al., 2019a).

Fourth, the regulation of CA4 does not rely on phytochrome superfamily members, as do the other known forms of CA. While the light color sensing system for CA4 remains to be discovered, it is clear that neither the genomes nor the CA4 GIs in the marine *Synechococcus* that are CA4-capable contain any genes encoding phytochrome superfamily members.

Thus far, the strength of CA4 research has been in its transdisciplinary nature, using a holistic approach that unifies our understanding of CA4 across scales, from the structural features of lyase and lyase-isomerase enzymes to the global ecological roles and evolution of CA4. Discoveries at each level of these investigations are being used to inform and advance our understanding at the other levels being studied. While much has been learned about the mechanisms, regulation, ecology, and evolution of CA4 in marine *Synechococcus*, much also remains to be learned. Over the past two decades, CA4 has been revealed as a process of global ecological importance whose study will enrich our understanding of light regulation, phenotypic plasticity, and the ecology of the marine environment.

FUNDING

This work was supported by National Science Foundation Grants (U.S.A.) MCB-1818187 and MCB-1029414 to D.M.K., MCB-1244339 and MCB-2017171 to W.M.S., MCB-2017274 to X.Y., and the French “Agence Nationale de la Recherche” programs EFFICACY (ANR-19-CE02-0019) and TAXCY (ANR-23-CE2-0007 as well as the European Union program Assemble+ (Horizon 2020, under grant agreement number 287589) to F.P. and L.G.

REFERENCES

- Anderson, L. K. & Toole, C. M. 1998. A model for early events in the assembly pathway of cyanobacterial phycobilisomes. *Mol Microbiol*, 30, 467-474.
- Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. 2005. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol Rev*, 29, 231-262.
- Brahamsha, B. 1996. A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl Environ Microbiol*, 62, 1747-1751.
- Bryant, D. A. & Gisriel, C. J. 2024. The structural basis for light harvesting in organisms producing phycobiliproteins. *Plant Cell*, doi: 10.1093/plcell/koae126
- Campbell, D. 1996. Complementary chromatic adaptation alters photosynthetic strategies in the cyanobacterium *Calothrix*. *Microbiol*, 142, 1255-1263.
- Deng, G. P., Liu, F., Liu, X. W. & Zhao, J. D. 2012. Significant energy transfer from CpcG2 phycobilisomes to photosystem I in the cyanobacterium *Synechococcus* sp. PCC 7002 in the absence of ApcD-dependent state transitions. *FEBS Lett*, 586, 2342-2345.
- Diakoff, S. & Scheibe, J. 1973. Action spectra for chromatic adaptation in *Tolypothrix tenuis*. *Plant Physiol*, 51, 382-385.
- Doré, H., Farrant, G. K., Guyet, U., Haguait, J., Humily, F., Ratin, M., Pitt, F. D., Ostrowski, M., Six, C., Brillet-Guéguen, L., Hoebeke, M., Bisch, A., Le Corguillé, G., Corre, E., Labadie, K., Aury, J. M., Wincker, P., Choi, D. H., Noh, J. H., Eveillard, D., Scanlan, D. J., Partensky, F. & Garczarek, L. 2020. Evolutionary mechanisms of long-term genome diversification associated with niche partitioning in marine picocyanobacteria. *Front Microbiol*, 11.
- Dufour, L., Garczarek, L., Gouriou, B., Clairet, J., Ratin, M. & Partensky, F. 2024. Differential acclimation kinetics of the two forms of type IV chromatic acclimators occurring in marine *Synechococcus* cyanobacteria. *Front Microbiol*, 15.
- Dufresne, A., Ostrowski, M., Scanlan, D., Garczarek, L., Mazard, S., Palenik, B., et al. 2008. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol*, 9, (5):R90.
- Dutkiewicz, S., Hickman, A.E., Jahn, O., Gregg, W.W., Mouw, C.B., Follows, M.J. 2015. Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeochem*, 12, 4447-4481.
- Engelmann, T. W. 1902. Untersuchungen über die qualitativen beziehungen zwieschen absorbtion des lichtetes und assimilation in pflanzenzellen. I. Das mikrospectrophotometer, ein apparat zur qualitativen mikrospectralanalyse. II. Experimentelle grundlangen zur ermittlung der quantitativen beziehungen zwieschen assimilationsenergie und absorbtiongrosse. III. Bestimmung der vertheilung der energie im spectrum von sonnenlicht mittels bacterienmethode und quantitativen mikrospectralanalyse. *Bot Z*, 42, 81-105.
- Everroad, C., Six, C., Partensky, F., Thomas, J. C., Holtendorff, J. & Wood, A. M. 2006. Biochemical bases of type IV chromatic adaptation in marine *Synechococcus* spp. *J Bacteriol*, 188, 3345-3356.
- Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincon, J., Zabala, L. L., Jiao, N. a. Z., Karl, D. M., Li, W. K. W., Lomas, M. W., Veneziano, D., Vera, C. S., Vrugt, J. A. & Martiny, A. C. 2013. Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc Natl Acad Sci USA*, 110, 9824-9829.
- Fujita, Y. & Hattori, A. 1960. Effect of chromatic lights on phycobilin formation in a blue-green alga, *Tolypothrix tenuis*. *Plant Cell Physiol*, 1, 293-303.
- Fujita, Y. & Hattori, A. 1962. Photochemical interconversion between precursors of phycobilin chromoproteids in *Tolypothrix tenuis*. *Plant Cell Physiol*, 3, 209-220.
- Gaidukov, N. 1902. Über den einfluss farbigen lichtetes auf die farbung lebender Oscillatorien. *Abh Preuss Akad Wiss*, V, 1-36.

- Gaidukov, N. 1903. Die farbervonderung bei den prozessen der komplementoren chromatischen adaptation. *Ber Dtsch Bot Ges*, 21, 517-522.
- Glazer, A. N. 1989. Light guides. Directional energy transfer in a photosynthetic antenna. *J Biol Chem*, 264, 1-4.
- Grébert, T., Doré, H., Partensky, F., Farrant, G. K., Boss, E. S., Picheral, M., Guidi, L., Pesant, S., Scanlan, D. J., Wincker, P., Acinas, S. G., Kehoe, D. M. & Garczarek, L. 2018. Light color acclimation is a key process in the global ocean distribution of *Synechococcus* cyanobacteria. *Proc Natl Acad Sci USA*, 115, E2010-E2019.
- Grébert, T., Garczarek, L., Daubin, V., Humily, F., Marie, D., Ratin, M., Devailly, A., Farrant, G. K., Mary, I., Mella-Flores, D., Tanguy, G., Labadie, K., Wincker, P., Kehoe, D. M. & Partensky, F. 2022. Diversity and evolution of pigment types in marine *Synechococcus* cyanobacteria. *Genome Biol Evol*, 14:evac035.
- Grébert, T., Nguyen, A. A., Pokhrel, S., Joseph, K. L., Ratin, M., Dufour, L., Chen, B., Haney, A. M., Karty, J. A., Trinidad, J. C., Garczarek, L., Schluchter, W. M., Kehoe, D. M. & Partensky, F. 2021. Molecular bases of an alternative dual-enzyme system for light color acclimation of marine *Synechococcus* cyanobacteria. *Proc Natl Acad Sci USA*, 118.
- Hackl, T., Laurenceau, R., Ankenbrand, M. J., Bliem, C., Cariani, Z., Thomas, E., Dooley, K. D., Arellano, A. A., Hogle, S. L., Berube, P., Leventhal, G. E., Luo, E. L. E., Eppley, J. M., Zayed, A. A., Beaulaurier, J., Stepanauskas, R., Sullivan, M. B., Delong, E. F., Biller, S. J. & Chisholm, S. W. 2023. Novel integrative elements and genomic plasticity in ocean ecosystems. *Cell*, 186, 47-62.e16.
- Haft, D. H., Basu, M. K. & Mitchell, D. A. 2010. Expansion of ribosomally produced natural products: a nitrile hydratase-and Nif11-related precursor family. *BMC Biology*, 8, 70.
- Hattori, A. & Fujita, Y. 1959. Effect of pre-illumination on the formation of phycobilin pigments in a blue-green alga, *Tolypothrix tenuis*. *J Biochem*, 46, 1259-1261.
- Haury, J. F. & Bogorad, L. 1977. Action spectra for phycobiliprotein synthesis in a chromatically adapting cyanophyte, *Fremyella diplosiphon*. *Plant Physiol*, 60, 835-839.
- Herdman, M., Castenholz, R.W., Waterbury, J.B. & Rippka, R. Form-genus XIII. *Synechococcus*. 2nd Ed. Boone D.R., Castenholz R.W., Editors. New York: Springer-Verlag; 2001.
- Hirose, Y., Chihong, S., Watanabe, M., Yonekawa, C., Murata, K., Ikeuchi, M. & Eki, T. 2019a. Diverse chromatic acclimation processes regulating phycoerythrocyanin and rod-shaped phycobilisome in cyanobacteria (vol 12, pg 715, 2019). *Mol Plant*, 12, 1167-1169.
- Hirose, Y., Shimada, T., Narikawa, R., Katayama, M. & Ikeuchi, M. 2008. Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein. *Proc Nat Acad Sci USA*, 105, 9528-9533.
- Hirose, Y., Song, C. H., Watanabe, M., Yonekawa, C., Murata, K., Ikeuchi, M. & Eki, T. 2019b. Diverse chromatic acclimation processes regulating phycoerythrocyanin and rod-shaped phycobilisome in cyanobacteria. *Mol Plant*, 12, 715-725.
- Holtrop, T., Huisman, J., Stomp, M., Biersteker, L., Aerts, J., Grébert, T., Partensky, F., Garczarek, L. & Van Der Woerd, H. J. 2021. Vibrational modes of water predict spectral niches for photosynthesis in lakes and oceans. *Nat Ecol Evol*, 5.
- Humily, F., Farrant, G. K., Marie, D., Partensky, F., Mazard, S., Perennou, M., Labadie, K., Aury, J. M., Wincker, P., Segui, A. N., Scanlan, D. J. & Garczarek, L. 2014. Development of a targeted metagenomic approach to study a genomic region involved in light harvesting in marine *Synechococcus*. *FEMS Microbiol Ecol*, 88, 231-249.
- Humily, F., Partensky, F., Six, C., Farrant, G. K., Ratin, M., Marie, D. & Garczarek, L. 2013. A gene island with two possible configurations is involved in chromatic acclimation in marine *Synechococcus*. *PLoS One*, 8, e84459.

- Jacobson, M. R., Brigle, K. E., Bennett, L. T., Setterquist, R. A., Wilson, M. S., Cash, V. L., Beynon, J., Newton, W. E. & Dean, D. R. 1989. Physical and genetic map of the major *nif* gene cluster from *Azotobacter vinelandii*. *J Bacteriol*, 171, 1017-1027.
- Kondo, K., Geng, X. X., Katayama, M. & Ikeuchi, M. 2005. Distinct roles of CpcG1 and CpcG2 in phycobilisome assembly in the cyanobacterium *Synechocystis* sp PCC 6803. *Photosynth Res*, 84, 269-273.
- Kondo, K., Ochiai, Y., Katayama, M. & Ikeuchi, M. 2007. The membrane-associated CpcG2 phycobilisome in *Synechocystis*: A new photosystem I antenna. *Plant Physiol.*, 144, 1200-1210.
- Kumarapperuma, I., Joseph, K. L., Wang, C., Biju, L. M., Tom, I. P., Weaver, K. D., Grebert, T., Partensky, F., Schluchter, W. M. & Yang, X. J. 2022. Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase. *Structure*, 30, 564-574.e563.
- Liu, L. N., Chen, X. L., Zhang, Y. Z. & Zhou, B. C. 2005. Characterization, structure and function of linker polypeptides in phycobilisomes of cyanobacteria and red algae: An overview. *BBA Bioenergetics*, 1708, 133-142.
- MacColl, R. 1998. Cyanobacterial phycobilisomes. *J Struct Biol*, 124, 311-334.
- Ohki, K., Watanabe, M. & Fujita, Y. 1982. Action of near UV and blue light on the photocontrol of phycobiliprotein formation: A complementary chromatic adaptation. *Plant Cell Physiol*, 23, 651-656.
- Olson, R. J., Chisholm, S. W., Zettler, E. R. & Armbrust, E. V. 1990. Pigments, size, and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. *Limnol Oceanogr*, 35, 45-58.
- Ong, L. J. & Glazer, A. N. 1991. Phycoerythrins of marine unicellular cyanobacteria. I. Bilin types and locations and energy transfer pathways in *Synechococcus* spp. phycoerythrins. *J Biol Chem*, 266, 9515-9527.
- Ong, L. J., Glazer, A. N. & Waterbury, J. B. 1984. An unusual phycoerythrin from a marine cyanobacterium. *Science*, 224, 80-83.
- Palenik, B. 2001. Chromatic adaptation in marine *Synechococcus* strains. *Appl Environ Microbiol*, 67, 991-994.
- Sánchez-Baracaldo, P., Bianchini, G., Di Cesare, A., Callieri, C. & Christmas, N. a. M. 2019. Insights into the evolution of picocyanobacteria and phycoerythrin genes (*mpeBA* and *cpeBA*). *Front Microbiol*, 10.
- Sanfilippo, J. E., Garczarek, L., Partensky, F. & Kehoe, D. M. 2019a. Chromatic acclimation in cyanobacteria: A diverse and widespread process for optimizing photosynthesis. *Annu Rev Microbiol*, 73, 407-433.
- Sanfilippo, J. E., Nguyen, A. A., Garczarek, L., Karty, J. A., Pokhrel, S., Strnat, J. A., Partensky, F., Schluchter, W. M. & Kehoe, D. M. 2019b. Interplay between differentially expressed enzymes contributes to light color acclimation in marine *Synechococcus*. *Proc Natl Acad Sci USA*, 116, 6457-6462.
- Sanfilippo, J. E., Nguyen, A. A., Karty, J. A., Shukla, A., Schluchter, W. M., Garczarek, L., Partensky, F. & Kehoe, D. M. 2016. Self-regulating genomic island encoding tandem regulators confers chromatic acclimation to marine *Synechococcus*. *Proc Natl Acad Sci USA*, 113, 6077-6082.
- Scheer, H. & Zhao, K. H. 2008. Biliprotein maturation: the chromophore attachment. *Mol Microbiol*, 68, 263-276.
- Schluchter, W. M., Shen, G. Z., Alvey, R. M., Biswas, A., Saunee, N. A., Williams, S. R., Mille, C. A. & Bryant, D. A. 2010. Phycobiliprotein biosynthesis in cyanobacteria: Structure and function of enzymes involved in post-translational modification. *Recent Advances in Phototrophic Prokaryotes*, 675, 211-228.
- Shukla, A., Biswas, A., Blot, N., Partensky, F., Karty, J. A., Hammad, L. A., Garczarek, L., Gutu, A., Schluchter, W. M. & Kehoe, D. M. 2012. Phycoerythrin-specific bilin lyase-isomerase controls blue-green chromatic acclimation in marine *Synechococcus*. *Proc Natl Acad Sci USA*, 109, 20136-20141.
- Six, C., Thomas, J. C., Garczarek, L., Ostrowski, M., Dufresne, A., Blot, N., Scanlan, D. J. & Partensky, F. 2007. Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol*, 8, R259.

- Six, C., Thomas, J. C., Thion, L., Lemoine, Y., Zal, F. & Partensky, F. 2005. Two novel phycoerythrin-associated linker proteins in the marine cyanobacterium *Synechococcus* sp. strain WH8102. *J Bacteriol*, 187, 1685-1694.
- Stomp, M., Huisman, J., Stal, L. J. & Matthijs, H.C.P. 2007. Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. *ISME J*, 1, 271-282.
- Tandeau De Marsac, N. 1977. Occurrence and nature of chromatic adaptation in cyanobacteria. *J Bacteriol*, 130, 82-91.
- Tandeau De Marsac, N. & Cohen-Bazire, G. 1977. Molecular composition of cyanobacterial phycobilisomes. *Proc Natl Acad Sci USA*, 74, 1635-1639.
- Vogelmann, T. C. & Scheibe, J. 1978. Action spectra for chromatic adaptation in the blue-green alga *Fremyella diplosiphon*. *Planta*, 143, 233-239.
- Watanabe, M. & Ikeuchi, M. 2013. Phycobilisome: architecture of a light-harvesting supercomplex. *Photosynth Res*, 116, 265-276.
- Wiethaus, J., Busch, A. W. U., Dammeyer, T. & Frankenberg-Dinkel, N. 2010. Phycobiliproteins in *Prochlorococcus marinus*: Biosynthesis of pigments and their assembly into proteins. *Eur J Cell Biol*, 89, 1005-1010.
- Xia, X.M., Lee, P., Cheung, S., Lu, Y.H & Liu, H.B. 2020. Discovery of euryhaline phycoerythrobilin-containing *Synechococcus* and its mechanisms for adaptation to estuarine environments. *Msystems*, 5.
- Xia, X. M., Partensky, F., Garczarek, L., Suzuki, K., Guo, C., Cheung, S. Y. & Liu, H. B. 2017. Phylogeography and pigment type diversity of *Synechococcus* cyanobacteria in surface waters of the northwestern Pacific Ocean. *Environ Microbiol*, 19, 142-158.
- Yerrapragada, S., Shukla, A., Hallsworth-Pepin, K., Choi, K., Wollam, A., Clifton, S., Qin, X., Muzny, D., Raghuraman, S., Ashki, H., Uzman, A., Highlander, S. K., Fryszczyn, B. G., Fox, G. E., Tirumalai, M. R., Liu, Y., Kim, S., Kehoe, D. M. & Weinstock, G. M. 2015. Extreme sensory complexity encoded in the 10-megabase draft genome sequence of the chromatically acclimating cyanobacterium *Tolypothrix* sp. PCC 7601. *Genome Announcements*, 3.
- Zhao, C., Höppner, A., Xu, Q. Z., Gärtner, W., Scheer, H., Zhou, M. & Zhao, K. H. 2017. Structures and enzymatic mechanisms of phycobiliprotein lyases CpcE/F and PecE/F. *Proc Natl Acad Sci USA*, 114, 13170-13175.

LEGENDS TO FIGURES

Figure 1. (A, left half) The color phenotypes of plated *Fremyella diplosiphon* filaments grown in green (left) or red (right) and displaying Type 3 CA. (A, right half) The color phenotypes of liquid cultures of *Synechococcus* RS9916 cells grown in green light (left) or blue light (right) and displaying Type 4 CA. (B) Diagram of the structure of a hemidiscoidal PBS appressed to the thylakoid membrane and in association with a photosystem II reaction center (middle, dark and light green). Rod structures are shown from their sides and the core structure is shown from its end. (C) The ratios of the fluorescence excitation maxima of PUB and PEB at 495 nm and 545 nm, respectively ($\text{Ex}_{\text{C}_{495:545}}$), are approximately 0.6 when grown in green light (left) and approximately 1.6 when grown in blue light (right). AU, arbitrary units

Figure 2. Diagram of genes in the CA4-A (upper) and CA4-B (lower) genomic islands found in 9916 and A15-62, respectively. Thin black arrows denote the proposed directions of action of the gene products (demonstrated for CA4-A, assumed for CA4-B). One kbp size marker is shown.

Figure 3. The MpeQ structure and proposed reaction mechanisms of bilin lyases and lyase-isomerases. (A) The single chain bilin lyase MpeQ adopts an all-helical structure. Two domains (E and F) corresponding to the E and F subunits of the heterodimeric CpcE/F form a question-mark architecture where the active site chamber is enclosed by the N-terminal extension (gray). The conserved residue Tyr318 (rendered in yellow spheres) is shown to mark the active site of MpeQ. (B) In a zoom-in view of the active site, two substrates (MpeA-Cys83 and PEB) are modeled relative to Tyr318 and Val319. We propose that a direct steric clash between the A-ring of PEB with Val319 in the Tyr318-mediated ligation reaction plays a critical role in shifting the relative population of products from MpeA-PEB (magenta) to MpeA-PUB (yellow).

Figure 4. Distribution of *Synechococcus* pigment types (PTs) (A) Relative abundance of each PT in the whole dataset (Total), at the surface, and at the deep chlorophyll maximum (DCM). (B) Map showing the global distribution of all *Synechococcus* PTs in surface waters along the Tara Oceans transect. Diameters of pies are proportional to the number of *cpcBA* reads normalized by the sequencing effort. Stations with less than 30 *cpcBA* or *mpeBA* reads are indicated by open circles, and those with no *cpcBA* reads by black dots. Numbers next to pies correspond to Tara Oceans stations. (C) Correlation analysis between *Synechococcus* PTs and environmental parameters measured for all sampled depths along the Tara Oceans transect. The scale shows the degree of correlation (blue) or anticorrelation (red) between two variables. Nonsignificant correlations (adjusted P value > 0.05) are indicated by crosses. Number of observations for each environmental parameter is indicated at bottom. Abbreviations: BAC, beam attenuation coefficient; Backscatt., backscattering; cDOM fluo., colored dissolved organic matter fluorescence; DCM, deep chlorophyll maximum; $\text{Irr}_{495:545}$, ratio of downwelling irradiance at 495 and 545 nm; IS, in situ; MLD, mixed layer depth; NPP, net primary production; PAR, photosynthetically active radiation; part., particulate; Φ_{sat} , satellite-based nonphotochemical quenching (NPQ)-corrected quantum yield of fluorescence (proxy for iron limitation). Modified from (Grébert et al., 2018).

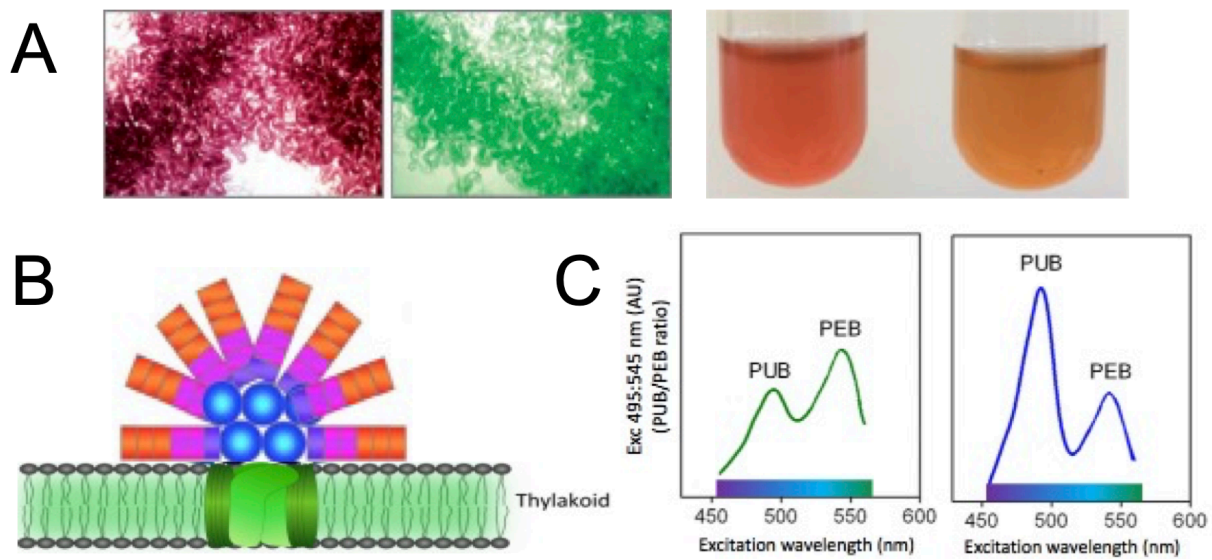


Figure 1

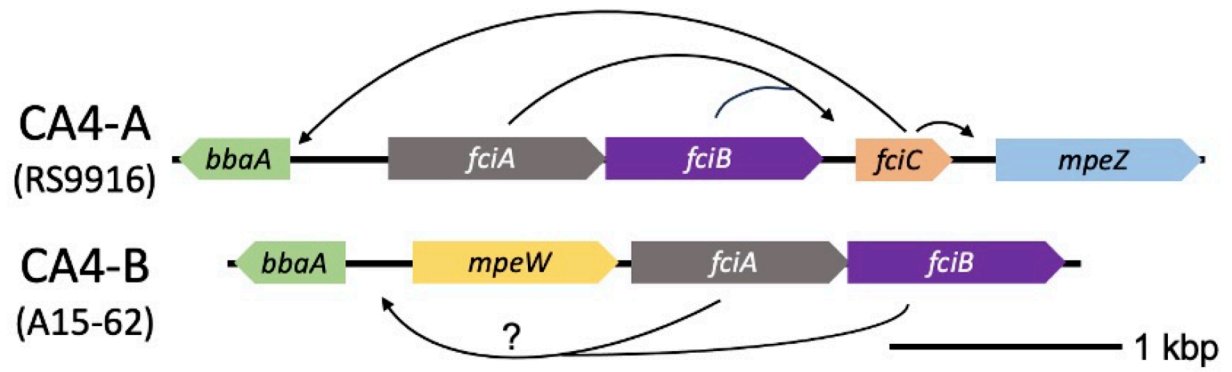


Figure 2

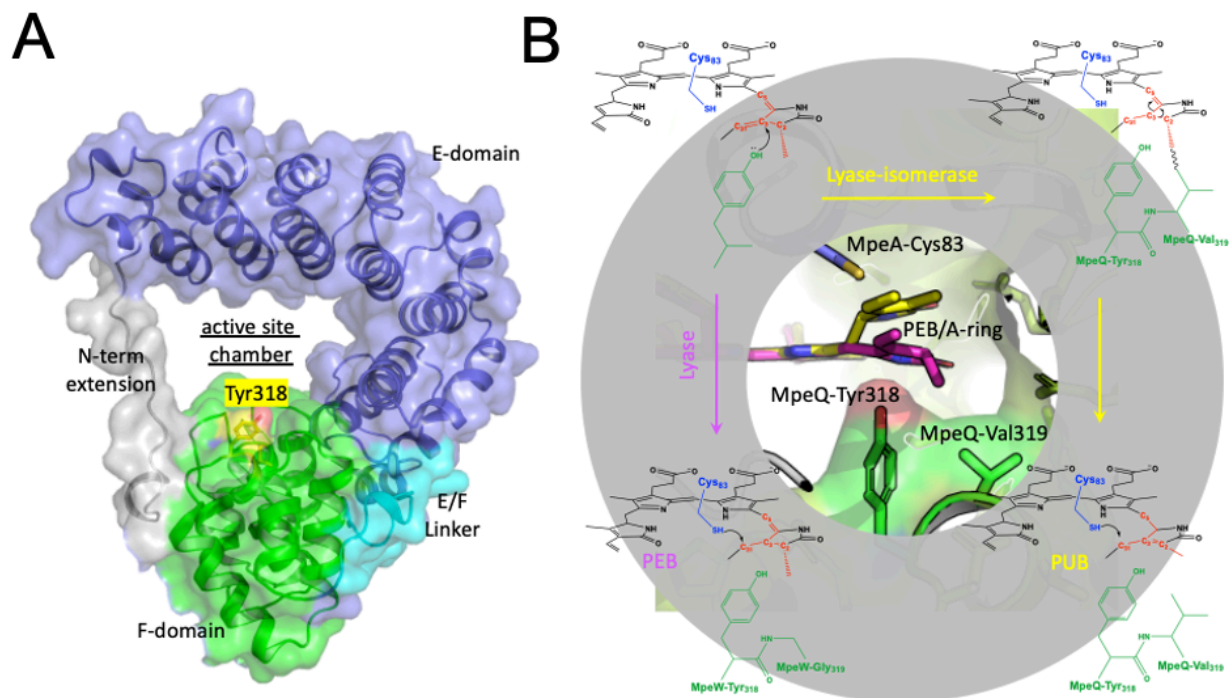


Figure 3

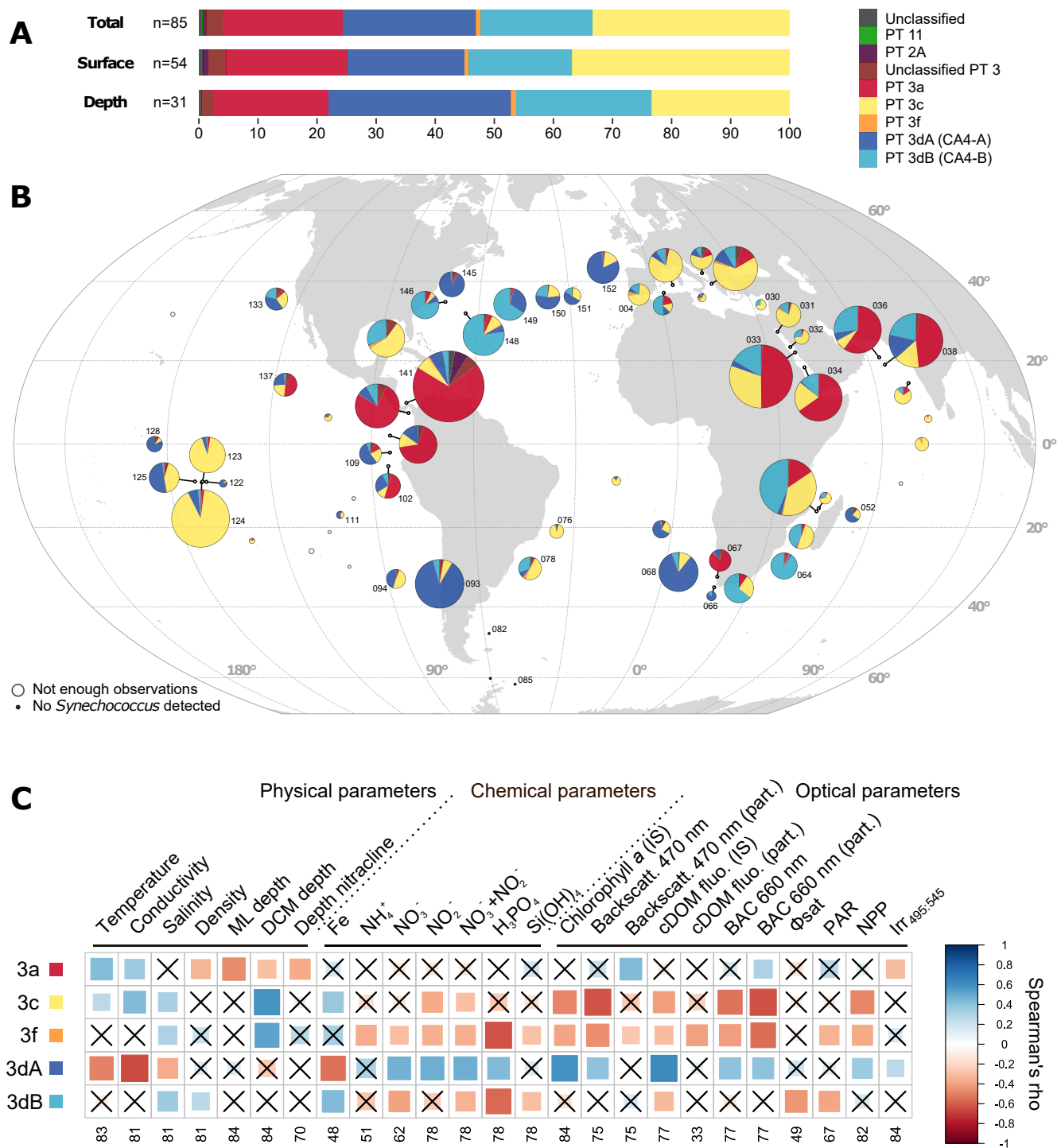


Figure 4