ENVIRONMENTAL DETERMINATION AND DYNAMIC REGULATION OF THE CARBON CONCENTRATING MECHANISM IN *FREMYELLA DIPLOSIPHON*

By

Brandon Rohnke

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biochemistry and Molecular Biology – Doctor of Philosophy

2019

PUBLIC ABSTRACT

ENVIRONMENTAL DETERMINATION AND DYNAMIC REGULATION OF THE CARBON CONCENTRATING MECHANISM IN *FREMYELLA DIPLOSIPHON*

By

Brandon Rohnke

Cyanobacteria are single-celled organisms that perform oxygenic photosynthesis. They are major contributors to Earth's carbon cycle and are a rich source of biological energy. Since they inhabit many diverse habitats, they must regulate photosynthesis based on resource availability; important resources for photosynthesis include light and carbon dioxide (CO_2) levels. One feature that cyanobacteria use to maximize photosynthetic efficiency involves the concentration of carbon dioxide around the carbon-fixing rubisco enzyme by trapping it in sub-cellular compartments, which allows the organisms to grow even at low CO_2 levels. Investigations in this study explore how a specific cyanobacterium, *Fremyella diplosiphon*, controls the carbon concentrating mechanism (CCM) in response to the color and intensity of light and in response to CO_2 levels. I present evidence that the ability of *F. diplosiphon* to detect light color and light intensity is involved in the regulation of the CCM. Furthermore, I contribute to the study of an underexplored protein that interacts with rubisco and with a possible role in responding to environmental conditions. I then analyze how the response of *F. diplosiphon* to resource availability impacts its ability to fix carbon. These investigations provide a picture of the many factors that contribute to the regulation of photosynthesis in cyanobacteria, with insight into the multiple components that could be leveraged for improved biofuel and biotechnology production.

ABSTRACT

ENVIRONMENTAL DETERMINATION AND DYNAMIC REGULATION OF THE CARBON CONCENTRATING MECHANISM IN *FREMYELLA DIPLOSIPHON*

By

Brandon Rohnke

Cyanobacteria are single-celled photoautotrophic organisms that are major contributors to global carbon fixation. Since the accumulation of significant amounts of oxygen (O_2) in the atmosphere, they have contended with decreased amounts of available inorganic carbon (C_i) with which to supply photosynthesis. This has led to the evolution of a carbon concentrating mechanism (CCM) that involves increasing the C_i uptake into the cell and utilizing a specialized bacterial microcompartment (BMC) called the carboxysome to trap carbon dioxide (CO_2) around the carbon-fixing enzyme, rubisco. Since the CCM is an integral part of cyanobacterial photosynthesis, it is expected to be regulated by many environmental factors that affect photosynthesis. Here, I present findings from studies on *Fremyella diplosiphon*, a model cyanobacterium that exhibits complementary chromatic acclimation (CCA), wherein photosynthetic pigmentation and efficiency are tuned in response to environmental light cues. A photoreceptor, RcaE, that controls CCA was shown to be important for determining the abundance and stoichiometry of CCM components. The size and abundance of carboxysomes were found to correlate with the ratio of carboxysomal cargo:shell in a Δ*rcaE* mutant strain, suggesting a role for RcaE in regulating carboxysome morphology. Additionally, *F. diplosiphon* is one of many cyanobacteria that express an activase-like cyanobacterial (ALC) protein, a homologue to rubisco activase that is an essential protein in plants for rubisco activity. The ALC protein is not coded for in the genomes of many model cyanobacteria but was predicted to be targeted to the carboxysome when present, thus representing an important factor in the nuanced

regulation of the CCM. Through contributions to a study highlighting the carboxysomal localization and enzymatic activity of the ALC protein, I provide evidence that ALC is involved in the cellular response to C_i availability. Additionally, computational modeling of the interaction between ALC and rubisco contrasted two potential binding sites and suggested that the interaction could depend on species of origin and post-transcriptional modification. Given the myriad of factors impacting carboxysome regulation, I then analyzed the carbon fixation capabilities of *F. diplosiphon* strains under variations in light quality, light quantity, and C_i availability. Assessment of carbon fixation behavior utilizing a novel application of carbon response curves to cyanobacteria was consistent with expectations and provided additional insight into which components of the CCM respond to environmental cues. The Δ*rcaE* mutant exhibited a noteworthy green-light-dependent limitation in carbon fixation, and further analyses suggested that this depended on rubisco levels and the expression of C_i-uptake genes. These findings, alongside the behavior of other cultivation conditions and *F. diplosiphon* strains, were used to distinguish at a preliminary level the components of carbon response curves in cyanobacteria as a method to assess carbon fixation behavior and the functional impacts of CCM regulation.

To Mom, for her love and support throughout this journey.
You mean the world to me.

To Mr. Brian Cox, in whose biochemistry class
I fell in love with the beauty of photosynthesis.

and

In loving memory of Will Classen, my Opa,
who encouraged me to stay curious through research.

ACKNOWLEDGEMENTS

I would first like to express my deep thankfulness to Dr. Beronda Montgomery, my amazing mentor who lives up to the word ‘mentor’ more than I could have ever hoped for. She is a blessing to so many people and does so with such kindness and enthusiasm that it is contagious. I have grown as a scientist and as a person under her care, and I hope to show the world what a treasure we have in her. I am also truly grateful to Dr. Cheryl Kerfeld, who has assisted my work through exciting collaboration and expert discussion. Thanks goes out to Drs. Thomas Sharkey, Claire Vieille, and Kristin Parent, who have served on my committee and have encouraged strong scientific thought through their careful and helpful insight into my work. I am indebted to Alicia Withrow who taught me how to use a transmission electron microscope and supported the many projects with materials, troubleshooting advice, and a sense of wonder at the beauty of cells. I am grateful to Sigal Lechno-Yossef, Shailendra Singh, and Bagmi Pattanaik for their collaborations found in this dissertation. A huge shout out goes to the skilled and enthusiastic undergraduate students I had the opportunity to mentor, Tim Rapp and Kiara Rodriguez-Perez, who both performed amazing work that contributed to this project as well. I would also like to thank the entirety of the Plant Research Laboratory at Michigan State University, who foster a fun and scientific community full of promising discussion and exciting ideas. Lastly, I would like to express my thanks to fellow members of the Montgomery Lab past and present who I have crossed paths with.

This work was supported by the U.S. Department of Energy (Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, grant no. DE-FG02-91ER20021 to B.L.M.).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

A ..... Carbon assimilation, the net rate of CO_2 uptake per unit area

AA ..... Ascorbic Acid

AAA+ ..... ATPases Associated with diverse cellular Activities

ALC ..... Rubisco-Activase-Like protein of Cyanobacteria

Ava ..... *Anabaena variabilis*

BG11/HEPES ..... BG11 medium with 20 mM HEPES (pH 8.0)

BMC ..... Bacterial Microcompartment

BMC-H ..... Bacterial Microcompartment Hexameric shell protein

BMC-P ..... Bacterial Microcompartment Pentameric shell protein

BMC-T ..... Bacterial Microcompartment Trimeric (pseudo-hexamer) shell protein

CA ..... Carbonic Anhydrase

CCA ..... Complementary Chromatic Acclimation

CCM ..... Carbon Concentrating Mechanism

Chl*a* ..... Chlorophyll *a*

C_i ..... Inorganic Carbon

CI ..... Confidence Interval

CO_2 ..... Carbon dioxide

CO_2s ..... $[CO_2]$ in sample chamber

C_q ..... Quantification cycle for qPCR

CRC ..... Carbon Response Curve

DCBQ 2,6-dichloro-*p*-benzoquinone

DCF ..... 2’,7’-dichlorodihydrofluorescein

DCF-DA ..... 2’,7’-dichlorodihydrofluorescein diacetate

EDX ..... Energy Dispersive X-ray

*Fd*ALC ..... ALC from *F. diplosiphon*

*Fd*CcmM ..... CcmM from *F. diplosiphon*

FRET ..... Förster Resonance Energy Transfer

GL ..... Green Light

HCR ..... High Carbon Requiring

K_m ..... Michaelis constant

M ..... Equatorial binding position of SSLD to rubisco

M* ..... 30 kDa band tagged by α-CcmM antibodies

M35 ..... C-terminal fragment of CcmM, CcmM-35

M58 ..... Full-length CcmM, CcmM-58

MGL ..... Medium Green Light

MRL ..... Medium Red Light

MSA ..... Multiple Sequence Alignment

mT2 ..... mTurquoise2

O_2 ..... Molecular oxygen

OD_λ ..... Optical Density at λ (nm)

PL ..... Photosynthetic Lamellae

PPB ..... Polyphosphate Body

PSII ..... Photosystem II

Q ..... Light dosage, i.e. photosynthetic photon flux density

qPCR ..... Quantitative Real-Time Polymerase Chain Reaction

RbcL ..... Large subunit of rubisco

RbcS ..... Small subunit of rubisco

Rca ..... Rubisco Activase

RL ..... Red Light

ROS ..... Reactive Oxygen Species

Rubisco ..... Ribulose-1,5-biphosphate carboxylase/oxygenase

RuBP *D*-ribulose-1,5-biphosphate

SD ..... Standard Deviation

SE ..... Standard Error

SF33 ..... Short-filament wild-type pigmentation strain of *F. diplosiphon*

SSLD ..... Small Subunit-Like Domain

STEM ..... Scanning Transmission Electron Microscopy

*Syn*6301 ..... *Synechococcus elongatus* PCC 6301

*Syn*6803 ..... *Synechocystis* sp. PCC 6803

*Syn*7942 ..... *Synechococcus elongatus* PCC 7942

*Syn*CcmM ..... CcmM from *Synechococcus elongatus* PCC 7942

TEM ..... Transmission Electron Microscopy

V_{max} ..... Maximum rate

WL ..... White Light

WT ..... Wild-Type Strain

YFP ..... Yellow Fluorescent Protein

CHAPTER 1

Balancing Photosynthesis and the Carbon Concentrating Mechanism in Cyanobacteria

Photosynthesis is one of the core energy sources in the global ecosystem. In a series of energy transfers and organic reactions, sunlight is converted into chemical energy and stored in high-energy organic molecules. This is the foundation of the global carbon cycle and reflects the ultimate biological source of the energy found in both food and fuels. Understandably, improving photosynthesis in general is a major area of interest for biotechnology research. Moreover, many disciplines benefit from the study of how regulatory mechanisms enable successful photosynthesis through coordinating a network of inter-dependent cellular processes under dynamic environmental conditions.

The main carbon-fixing enzyme ribulose-1,5-biphosphate carboxylase/oxygenase (rubisco) is a component of photosynthesis that is tightly regulated by many mechanisms (see reviews[18,84,145,194]). This enzyme has two substrates: CO_2 and a pentose sugar molecule, RuBP. Through the reactions of the Calvin-Benson cycle, RuBP is regenerated with a net incorporation of CO_2 into organic molecules[37]. On one hand, this reaction is dependent on how quickly RuBP can be recycled, which is in turn dependent on how much energy has been acquired from the light reactions of photosynthesis. However, the reaction also depends heavily on CO_2 availability, with atmospheric concentrations of CO_2 being rate limiting in many natural conditions[55,115]. Discerning between these effects through modeling has driven significant advances in understanding photosynthetic rates and parameters[55,115]. In addition to rate limiting factors, it is also important to consider why overall flux through the light reactions and Calvin-Benson cycle should be kept in balance; rubisco reflects a significant investment of nitrogen and energy stores so should not be in excess, but insufficient rubisco activity can lead to unused light energy causing lethal damage through photo-oxidative stress, such as the formation of reactive

oxygen species (ROS) (see review[166]). Overall, rubisco must be regulated by a substantial network of factors in photosynthetic organisms for efficient and safe photosynthesis.

Of the many organisms with the capacity to undergo photosynthesis, cyanobacteria are major contributors to global carbon fixation due to their abundance in diverse aquatic ecosystems[58,110]. Additionally, cyanobacteria hold a significant role in the evolutionary history of photosynthesis; they have an ancestral relationship with chloroplasts in green algae and higher order plants[45,62,225]; in addition, their dominance in the *Proterozoic* ecosphere drove atmospheric oxygenation[53,177,220]. Since O_2 can also react with rubisco in an energetically wasteful side-reaction (photorespiration), the rising O_2:CO_2 ratio posed an issue to cyanobacteria's continued ability to photosynthesize, driving the evolution of a carbon concentrating mechanism (CCM).

The CCM can be conceptually divided into two aspects: 1) expression of systems that increase inorganic carbon (C_i) uptake capabilities of the cell and 2) the encapsulation of rubisco into a compartment with a protein shell that serves as a barrier to CO_2 diffusion. Comprehensive research has established that active transport and CO_2 hydration are used to increase C_i uptake into the cell (see review[30,155]) as HCO_3^-, which is only appreciably converted to CO_2 in the presence of active carbonic anhydrase (CA). Both rubisco and CA were found to be compartmentalized in proteinaceous bacterial microcompartments (BMCs) called carboxysomes[60,91,210], first identified as polyhedral, electron-dense bodies during ultrastructural analysis of various cyanobacteria[59,82]. Unlike the cellular membrane, the protein shell of the carboxysome serves as a barrier to CO_2 diffusion[46]. Since the high levels of HCO_3^- accumulated in the cytosol can diffuse into the carboxysome but are trapped once converted to CO_2,

encapsulated rubisco experiences a locally high [CO_2], with consumption of CO_2 further driving cellular C_i uptake. Thanks to this CCM, cyanobacteria are able to grow at relatively low levels of available C_i despite cyanobacterial rubisco having a very high K_m for CO_2 compared to the rubisco of plants[11,86,204]. When the CCM is compromised, cyanobacteria can no longer grow at ambient [CO_2], exhibiting a high-carbon requiring (HCR) phenotype.

As with the C_i-uptake system, significant progress has been made in identifying the composition and biogenesis of the carboxysome. Two types of carboxysomes have been identified depending on the form of rubisco, with cyanobacteria containing Form IA rubisco having α-carboxysomes and Form IB rubisco having β-carboxysomes. Though their components are distinct and there are differences in assembly pathways, the two carboxysomes contain strong parallels in components[12,54,91,160,219]. In both, the functional unit of the protein shell is a protein domain (pfam00936) that forms six-sided hexameric structures (BMC-H) which can self-assemble into the faces of the carboxysome[56,92] (Figure 1.1A). Importantly, the hexameric structure defines a pore; for the major shell protein in β-carboxysomes, CcmK2, it is an ~7Å positively-charged pore[92] that is expected to allow the passage of small, negatively charged ions such as HCO_3^-[120]. However, the pore properties are variable in many paralogues of shell proteins[36,192], suggesting that shell composition can be used to adjust pore permeability[193], and thus associated carboxysome function. Proteins with two pfam00936 domains in tandem (BMC-T) that likely form stacked hexamers, and pentameric proteins (BMC-P) form the vertices (Figure 1.1A).

In β-carboxysomes, the assembly of the carboxysome depends on protein-protein interactions and begins with the aggregation of rubisco and CcmM[38,213]. CcmM proteins contain 3-5 C-

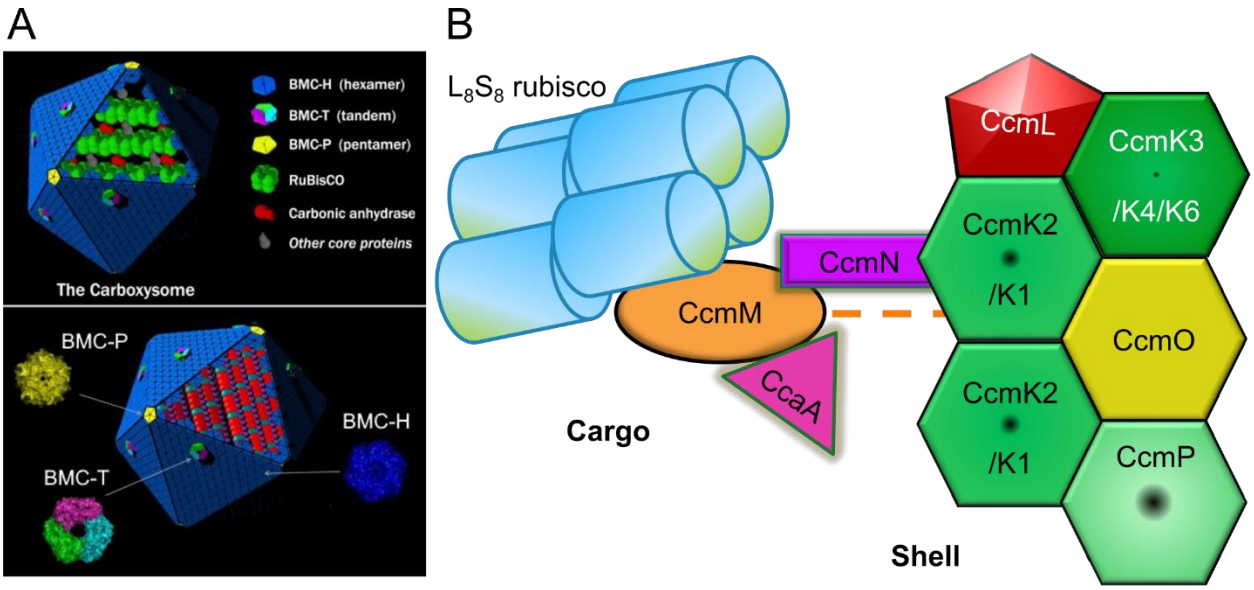


Figure 1.1: Protein-level schematics of β-carboxysomes. **(A)** Carboxysome structure modeled by Raul Gonzalez, Seth Axen, Markus Sutter, Sarah Newnham, Clément Aussignargues and Cheryl Kerfeld, reproduced under a Creative Commons Attribution Share Alike 4.0 International License. Images has been modified to remove protein side views and to enlarge text. **(B)** Schematic of known interactions between β-carboxysome proteins. CcmK2 is the major shell protein with a close paralogue CcmK1; CcmK3, K4, & K6 are additional paralogs that can form homodimers, heterodimers, & stacked dodecamers[192,193], CcmO & CcmP are BMC-T shell proteins, and CcmL is a BMC-P protein. Dark spots indicate approximate relative pore size based on available crystal structures.

terminal repeats of a domain with homology to RbcS[116,156] and the multiple repeats contribute to a paracrystalline network of rubisco in the cargo of carboxysomes[112,113,171,213]. The N-terminal domain of CcmM either acts as a γ-class CA itself[6], or interacts with cyanobacterial β-class CA CcaA[125], in addition to interactions with CcmN and CcmK2[44,94]. Shell protein interactions drive the encapsulation of the cargo with CcmL serving to cap the vertices[38,159] (Figure 1.1B). This self-assembly process and the morphology of carboxysomes likely depends on the stoichiometry of carboxysome components[113,114,200], highlighting one area with potential for regulation that can also provide insight into control of BMCs in biotechnology.

Given the tight and multi-faceted regulation of rubisco in plants, it is expected that the CCM, which enhances rubisco activity in cyanobacteria, is regulated at many levels by environmental cues and photosynthetic rates. The CCM reflects a significant metabolic investment, which is

evidenced by the inducible characteristics of many CCM components in response to high light[61,76,79,123,199] or low C_i availability[31,48,124,214], conditions which increase the need for available C_i. In stable conditions, these large changes occur rapidly (likely within minutes) in the C_i-uptake system and more slowly (on the order of days) for carboxysomes. With the importance for balancing carbon fixation rate and flux through the light reactions in dynamic conditions, cyanobacteria are expected to have nuanced methods to fine-tune factors such as carboxysome morphology, abundance, shell composition, and packing density in response to environmental cues. *Fremyella diplosiphon* is a freshwater, filamentous cyanobacteria that is notable for its characteristic acclimation to external light cues. This responses is known as complementary chromatic acclimation (CCA) and includes distinct pigmentation phenotypes that allow acclimated cyanobacteria to thrive in different light conditions, such as when transitioning to deeper sections of the water column where red light becomes less available[129]. In this dissertation, I highlight contributions made toward identifying the impact of a light-sensing photoreceptor responsible for CCA on the regulation of the stoichiometry of CCM components, exploring the potential regulatory role of a rubisco activase homologue in cyanobacteria, and examining the functional impacts of CCM regulation on carbon assimilation rates in the cyanobacterium *Fremyella diplosiphon*.

CHAPTER 2

RcaE-Dependent Regulation of Carboxysome Structural Proteins has a Central Role in Environmental Determination of Carboxysome Morphology and Abundance in *Fremyella diplosiphon*

This chapter contains information published in:

Rohnke, B. A., Singh, S. P., Pattanaik, B., and Montgomery, B. L. (2018) RcaE-dependent regulation of carboxysome structural proteins has a central role in environmental determination of carboxysome morphology and abundance in *Fremyella diplosiphon*. *mSphere 3*, e00617-17. doi: 10.1128/mSphere.00617-17

2.1 Abstract

Carboxysomes are central to the CCM and carbon fixation in cyanobacteria. Although the structure is well understood, roles of environmental cues in the synthesis, positioning, and functional tuning of carboxysomes have not been systematically studied. *F. diplosiphon* is a model cyanobacterium for assessing impacts of environmental light cues on photosynthetic pigmentation and tuning of photosynthetic efficiency during CCA, which is controlled by photoreceptor RcaE. Given the central role of carboxysomes in photosynthesis, we investigated roles of light-dependent RcaE signaling on carboxysome structure and function. A Δ*rcaE* mutant exhibits altered carboxysome size and number, *ccm* gene expression, and carboxysome protein accumulation relative to wild type (WT). Several Ccm proteins, including carboxysome shell proteins and core-nucleating factors, overaccumulate in Δ*rcaE* cells relative to WT. Additionally, levels of carboxysome cargo rubisco are lower to unchanged in Δ*rcaE* compared to WT. This shift in ratios of carboxysome shell and nucleating components to the carboxysome cargo appears to drive carboxysome morphology and abundance dynamics. Carboxysomes are also occasionally mislocalized spatially to the periphery of spherical mutants within thylakoid membranes suggesting that carboxysome positioning is impacted by cell shape. The RcaE photoreceptor links perception of external light cues to regulating carboxysome structure and function, and thus, cellular capacity for carbon fixation.

2.2 Importance

Carboxysomes are proteinaceous subcellular compartments, or bacterial organelles, found in cyanobacteria that consist of a protein shell surrounding a core primarily composed of the enzyme rubisco that is central to the CCM and carbon fixation. Whereas significant insights have

been gained regarding the structure and synthesis of carboxysomes, limited attention has been given to how the size, abundance and protein composition is regulated to ensure optimal carbon fixation in dynamic environments. Given the centrality of carboxysomes in photosynthesis, we provide an analysis of the role of a photoreceptor RcaE, which functions in matching photosynthetic pigmentation to the external environment during CCA and thereby optimizing photosynthetic efficiency, in regulating carboxysome dynamics. Our data highlight a role for RcaE in perceiving external light cues and regulating carboxysome structure and function and, thus, cellular capacity for carbon fixation and organismal fitness.

2.3 Introduction

Some cyanobacterial strains tune photosynthetic capacity to environmental cues, including changes in the availability of light. *F. diplosiphon* is a filamentous, freshwater cyanobacterium that exhibits CCA, which is a process to tune photosynthetic pigment type and levels to changes in the prevalent wavelengths of external light[129]. In *F. diplosiphon*, CCA-associated changes occur primarily in response to the presence and abundance of red versus green wavelengths of light[19]. In red-enriched light, *F. diplosiphon* accumulates red-absorbing, green-colored phycocyanin in light-harvesting complexes to maximize light absorption for photosynthesis. Conversely, under green-enriched conditions, *F. diplosiphon* accumulates green-absorbing, red-colored phycoerythrin for promoting light harvesting. In addition to pigmentation changes, cell shape and filament length also are controlled by light during CCA[19]. Cyanobacteriochrome (phytochrome-related) photoreceptor RcaE controls both the light-dependent regulation of pigmentation[89,209] and cell and filament morphologies[23,189] characteristic of CCA in *F. diplosiphon*. As wavelength-dependent tuning of pigmentation is linked to the maintenance of

optimal photosynthetic efficiency[39], RcaE has a role in tuning photosynthetic potential to external light cues. In prior studies, we noted a reduction in growth of a Δ*rcaE* mutant strain relative to WT under ambient air in red light (RL) or green light (GL)[187], and that the expression of genes associated with C_i-uptake were generally upregulated in the Δ*rcaE* mutant relative to WT[130]. Together, these phenotypes suggest an HCR phenotype associated with defects in bicarbonate uptake, with C_i-uptake, or with some part of the CCM.

Apart from suggesting a potential state of C_i deficiency in Δ*rcaE* cells, the impact of an absence of RcaE on the expression of C_i-uptake genes is particularly significant as light has previously been reported to be required under low C_i conditions for the expression of genes impacting inducible C_i-uptake systems of the CCM in *Synechocystis* sp. strain PCC 6803 (hereafter *Syn*6803)[123,124]. Because light is required for this process, redox or phytochrome signals were implicated in the light-dependent cellular response to low C_i in cyanobacteria[155]. The regulation of inorganic carbon uptake genes involved in the CCM in a Δ*rcaE* mutant provided genetic evidence of involvement of the photoreceptor RcaE in responses to low C_i[130].

The CCM is modular with distinct components consisting of the C_i-uptake systems at the membrane in addition to the intracellular carboxysome subcompartment[14,30,86]. The carboxysome is a specialized BMC containing rubisco, which functions in carbon fixation[60,91]. Although there have been significant insights into the structural make-up of carboxysomes[91] and assembly principles of BMC shells[201], there have been limited insights about the environmental inputs that regulate the synthesis, positioning, and potential functional tuning of carboxysomes. Prior studies demonstrated that regular distribution and positioning of carboxysomes along the long axis of the

cell is critical for maintaining carboxysome partitioning and associated cellular fitness during cell division[174]. Notably, carboxysomes increase in number in low carbon conditions in WT *Syn*6803[48] and *Synechococcus* strains[127,199,219]. Additionally, increased light intensity leads to an increase in the transcription of carboxysome genes[76,79] and synthesis of carboxysomes in *Synechococcus elongatus* PCC 7942 (hereafter *Syn*7942)[199]. The increase in carboxysome number under elevated light conditions presumably increases carbon fixation capacity as a coordinate and long-term acclimation response to an increase in photosynthetic potential under increased availability of photons to drive electron transport. In addition to extended or continuous high light, the expression of carboxysome-related genes increases in the light cycle of diurnal conditions[61,217] or subjective day during circadian growth[81]. In one proteomic study with *Cyanothece*, carboxysome proteins also accumulated to higher levels in the light phase of a diel cycle[7]. The mechanisms regulating environmental regulation of carboxysomes in cells have not received significant experimental attention.

Here, we report on the investigation of the regulation of cellular responses to dynamic light conditions, including coordinate regulation of light absorption capabilities and carboxysome number, structure, and function in the CCA-capable *F. diplosiphon*. A Δ*rcaE* mutant that is incapable of normal regulation of CCA exhibits smaller and apparently more numerous carboxysomes than WT cells. Thus, we assessed *ccm* gene expression and protein accumulation in the red and green wavelengths critical for CCA and in WT compared to a Δ*rcaE* strain. Given the known phenotypes of altered cell shape[23] and high accumulation of reactive oxygen species (ROS) in the Δ*rcaE* strain[186], we also assessed whether cell shape and intracellular ROS levels have indirect impacts on carboxysome structure using cell-shape mutants and ROS-mitigating

compounds. Our results suggest a role for RcaE, including transcriptional regulation of *ccm* genes, in controlling carboxysome structure and number that may be linked to functional tuning of carboxysomes in response to external light cues.

2.4 Results

2.4.1 RcaE regulates carboxysome size and abundance in *F. diplosiphon*

RcaE is known to control both light-dependent regulation of pigmentation[89,209] and cell and filament morphologies[23] in *F. diplosiphon.* A RcaE-deficient strain of *F. diplosiphon* grows slower than WT in ambient air[187] (Figure 2.1) and has increased expression of C_i-uptake

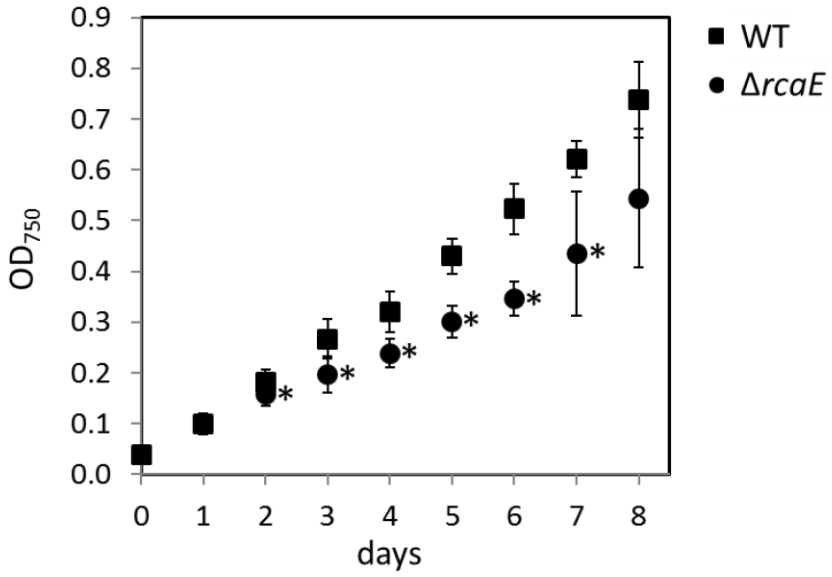


Figure 2.1: Growth curve of *F. diplosiphon* WT and Δ*rcaE* strains under WL. Growth rates of the WT and Δ*rcaE* strains in ambient air were estimated using OD_{750} measured once every 24 h for 7 d under continuous RL-enriched WL (~35-40 µmol m^{-2} s^{-1}). Data points represent averages (± SD). *, $p < 0.05$ (for results of comparisons between the Δ*rcaE* and WT strains determined using an unpaired *t* test).

genes[130], which together suggest that the Δ*rcaE* mutant has an HCR phenotype. To assess subcellular differences in the Δ*rcaE* mutant that may underlie such an HCR phenotype, we performed detailed ultrastructure analyses of WT and Δ*rcaE* strains grown under both RL and GL conditions using transmission electron microscopy (TEM). WT cells were more brick shaped

and elongated under GL than RL (Figure 2.2A), and Δ*rcaE* was spherical in both RL and GL (Figure 2.2A), as previously described for confocal laser scanning microscopy-based images[23]. Photosynthetic lamellae (PL) were regularly arranged around the cell perimeter in WT cells grown under both RL and GL (Figure 2.2A). By contrast, PL were more irregularly arranged or dispersed in the Δ*rcaE* mutant cells under GL or RL. Carboxysomes were larger in size in WT cells as compared to Δ*rcaE* mutant cells, independent of light conditions (Figure 2.2A & 2.2B; Table 2.1). Additionally, a comparison between RL and GL showed that carboxysomes were smaller under GL than RL in both WT and Δ*rcaE* strains. Although smaller in size, carboxysome number per cell section was significantly greater in Δ*rcaE* than WT under both RL and GL conditions (Figure 2.2C; Table 2.1).

2.4.2 RcaE regulates carboxysome-associated gene expression and protein accumulation in *F. diplosiphon*

Given the observed differences in carboxysome size, we assessed whether there were mutations in the sequences of known carboxysome genes by amplifying and sequencing targeted genomic regions. Similar to other cyanobacteria containing β carboxysomes, the key components of the carboxysome are encoded in a core *ccm* operon in *F. diplosiphon*, with additional shell proteins encoded in disparate satellite locations in the genome (Figure 2.3A). The core *ccm* operon encodes the shell proteins CcmK2, CcmK1, CcmL and CcmO, as well as other components, including CcmN which is essential for shell assembly and CcmM which facilitates rubisco nucleation (see review[91]). CcmP is encoded at a separate genomic location (Figure 2.3A). *F. diplosiphon* is one type of cyanobacterium that contains an expanded set of paralogs for proteins which comprise the carboxysome shell, including CcmK3 and CcmK4. These non-essential

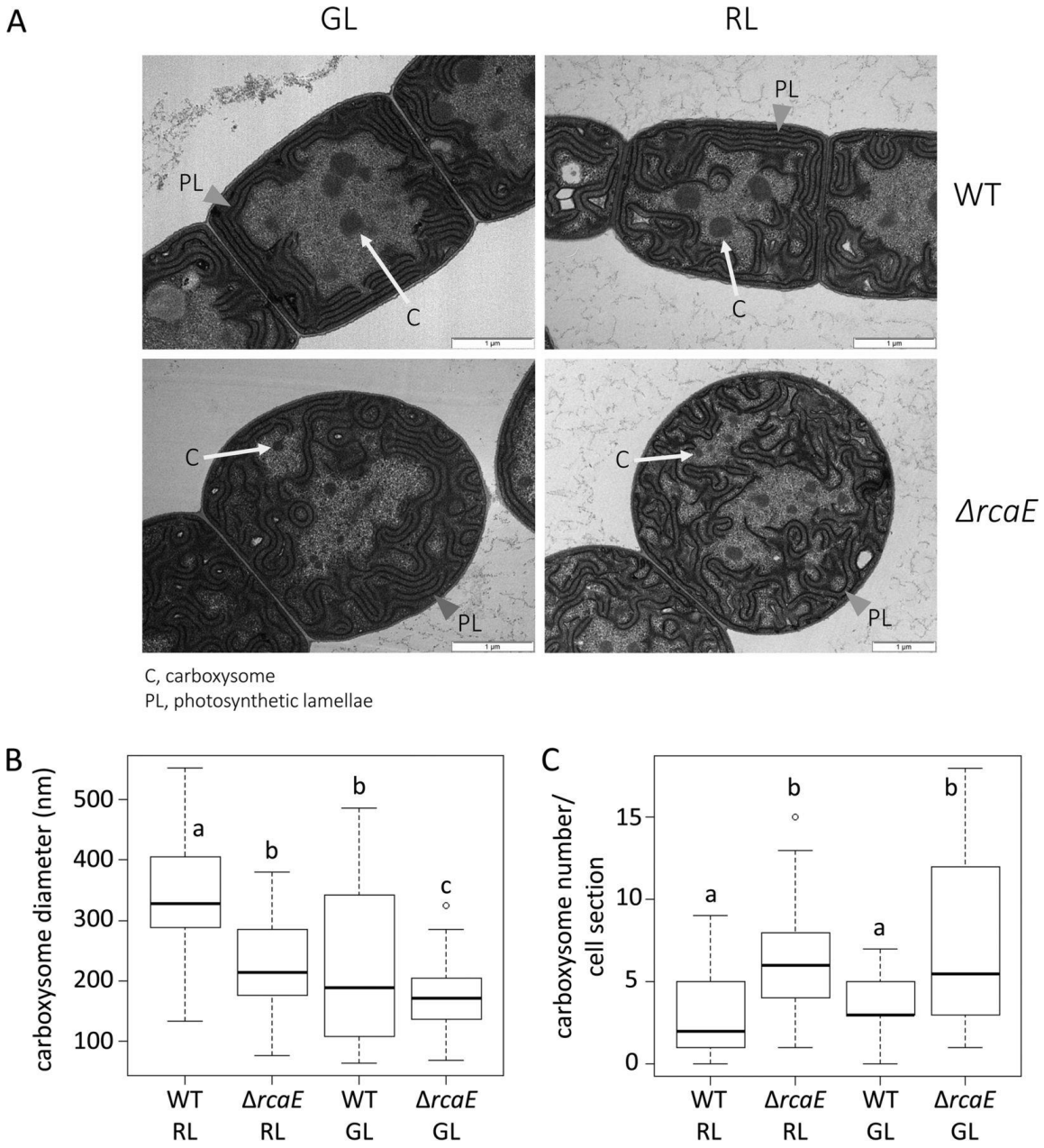


Figure 2.2: Carboxysome structure, size, and abundance determination in *F. diplosiphon* strains under GL and RL conditions. (A) TEM analysis of cellular morphology of *F. diplosiphon*. Representative images of WT pigmentation strain and Δ*rcaE* mutant strain under GL and RL. C, carboxysomes indicated by white arrows; PL, photosynthetic lamellae indicated by gray arrowheads. Bar represents 1 µm. Carboxysome **(B)** size and **(C)** number measurements of WT and Δ*rcaE* strains under GL and RL. To determine size, the maximum diameter of at least 25 carboxysomes were measured from each strain under each growth condition and presented as a boxplot. Boxplots were used as they present the entire data population spread, ordered from smallest to largest. The horizontal bold line inside each box corresponds to the median, and the box covers the 2nd and 3rd quartile groups (the middle 50% of all values). The vertical line below the box corresponds to the 1st quartile group (the smallest 25% of all values) and the line above the box corresponds to the 4th quartile group (the largest 25% of all values). Presenting the entire spread of data allows for visualization of differences between population spreads. Corresponding averages (± SE) can be found in Table 2.1. Statistical analyses were conducted using a Welch two sample *t* test performed in R. Identical letters over bars represent a homogenous mean group ($p > 0.05$), different symbols indicate a significant difference ($p < 0.05$) from others.

Table 2.1: Quantification of average carboxysome size and average carboxysome number per cell section in WT and Δ*rcaE* strains of *F. diplosiphon* from Figure 2.2.

	RL[a]		GL		MRL		MGL		RL + AA	
	WT	Δ*rcaE*[c]	WT	Δ*rcaE*	WT	Δ*rcaE*	WT	Δ*rcaE*	WT	Δ*rcaE*
Carboxysome size (nm)[b]	340 ± 19	224 ± 12 *	227 ± 19 #	174 ± 5 *,#	380 ± 22	317 ± 18 *,**	398 ± 28 **	250 ± 13 *,**,#	328 ± 14	214 ± 11 *
Carboxysomes/cell section	3.0 ± 0.3	6.2 ± 0.3 *	3.8 ± 0.2	7.2 ± 0.3 *	3.5 ± 0.2	7.1 ± 0.3 *	2.8 ± 0.2 **	6.7 ± 0.3 *	3.5 ± 0.2	6.4 ± 0.2 *
Sample size (n) for carboxysome size measurements	27	43	45	106	35	71	28	47	29	51
Sample size (n) for carboxysomes/ cell section measurements	91	186	114	215	105	178	28	47	35	64

[a] Indicates light conditions under which WT and Δ*rcaE* cells are grown, i.e., RL, red light at ~10 to 12 µmol m^{-2} s^{-1}; GL, green light at ~10 to 12 µmol m^{-2} s^{-1}; MRL, medium red light at ~30 µmol m^{-2} s^{-1}; MGL, medium green light at ~30 µmol m^{-2} s^{-1}; RL + AA, red light at ~10 to 12 µmol m^{-2} s^{-1} with added ascorbic acid (AA) at 2 mM.
[b] Numbers for carboxysome size and carboxysome/cell section are represented as average ± SE.
[c] Statistical analyses, $p < 0.05$ indicated as follows: *, WT vs. Δ*rcaE* in same condition; **, low light vs. medium light for same light quality for the same strain; #, GL vs. RL for same light intensity for the same strain.

paralogs are often at disparate locations from the core *ccm* operon and provide an expanded set of carboxysome shell subunits that have been hypothesized to afford selective advantages by altering carboxysome shell permeability, and thus function, under dynamic growth conditions[159,192]. Based on Sanger sequencing performed on PCR-amplified, *ccm* gene-containing regions of the genome, we identified no mutations in the sequences of known *ccm* or carboxysome genes (data not shown) in the Δ*rcaE* strain relative to WT.

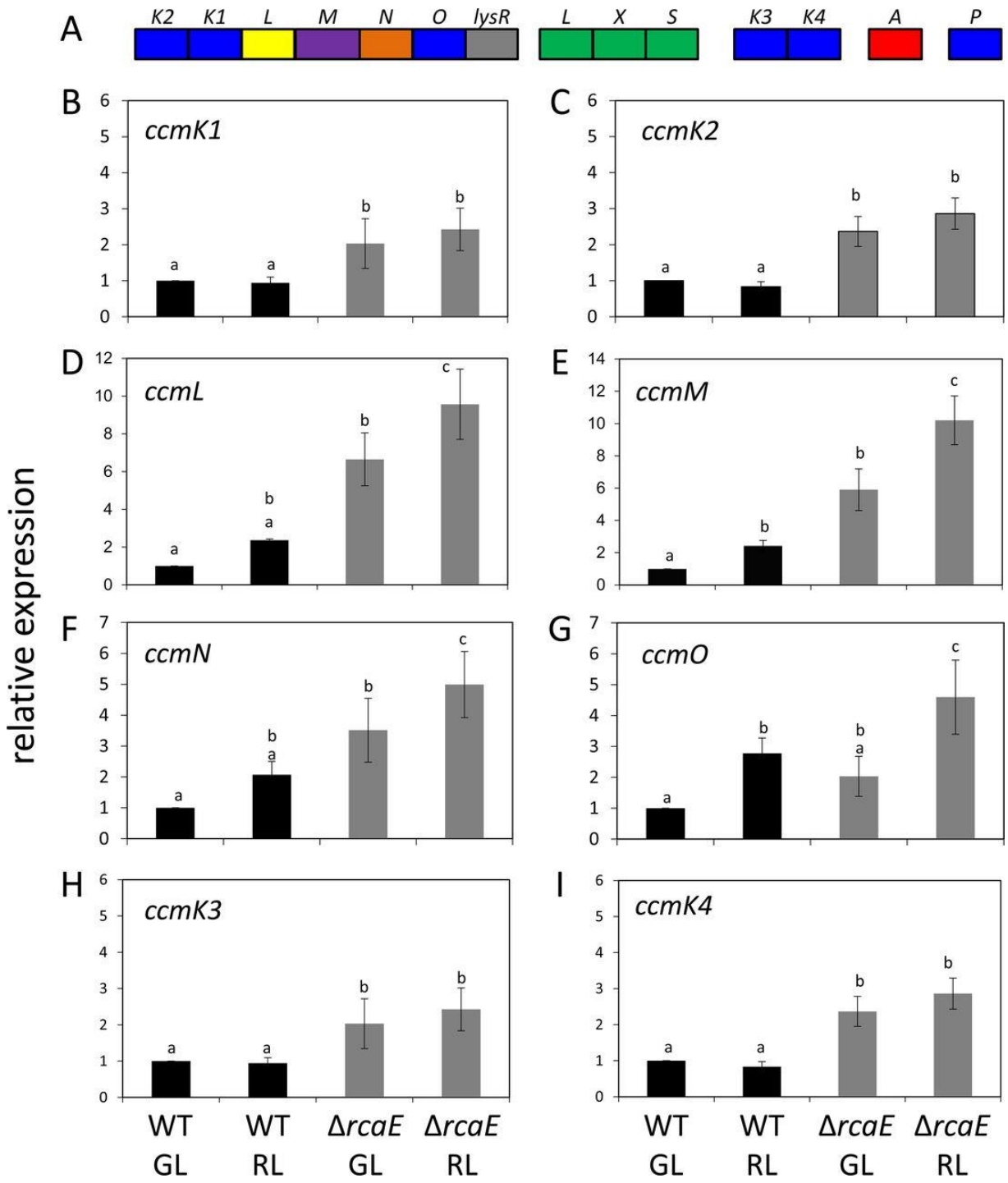


Figure 2.3: Carboxysome operons and qPCR-based gene expression analyses in *F. diplosiphon*. (A) Carboxysome-associated genes and operons found in *F. diplosiphon*. **(B** to **I)** Data represent levels of expression of *ccm* genes, including **(B)** *ccmK1*, **(C)** *ccmK2*, **(D)** *ccmL*, **(E)** *ccmM*, **(F)** *ccmN*, **(G)** *ccmO*, **(H)** *ccmK3*, and **(I)** *ccmK4*, in WT and Δ*rcaE* strains grown under GL or RL. Levels of expression of genes are presented relative to the results determined for the internal control *orf10B*, and the data in each panel are shown relative to the expression level of the gene of interest in WT cells in GL. Bars represent averages (± SD) of data from three independent biological replicates. Identical letters over bars represent homogenous mean groups ($p > 0.05$); different symbols indicate a statistically significant difference ($p < 0.05$) from others.

We proceeded to assess differences in the expression of *ccm* genes using data from a prior RNA-sequencing analysis comparing WT and Δ*rcaE* strains[146]. All *ccm* genes, with the exception of *ccmO*, *ccmK3*, and *ccmP*, exhibited significantly increased mRNA levels in the Δ*rcaE* mutant relative to WT in RNA-seq analysis (Table 2.2). Notably, *rbc* genes were largely downregulated in the Δ*rcaE* mutant compared to WT (Table 2.2). We confirmed differences for select *ccm* genes by quantitative real-time polymerase chain reaction (qPCR) (Figure 2.3B).

To assess whether the observed transcriptional responses were also apparent at the protein level, CcmM, CcmK2, and RbcL proteins were examined using immunoblot analyses. CcmM and CcmK2 proteins accumulated to higher levels in Δ*rcaE* relative to WT (Figure 2.4), reflecting that these factors are regulated at the transcriptional level. CcmM exhibits two forms in cells, due to an internal ribosome binding site on the transcript[114]. This results in two distinct forms of CcmM, i.e., ~58 kDa (CcmM-58 or M58) and ~35 kDa (CcmM-35 or M35), that accumulate in cyanobacteria. CcmM-58 levels especially were elevated in Δ*rcaE* relative to WT, although CcmM-35 levels were somewhat elevated particularly in GL (Figure 2.4A). Thus, the ratio of CcmM-58 to CcmM-35 changed in WT relative to the Δ*rcaE* mutant, in addition to total CcmM levels. Prior analyses in which the ratio of M58:M35 was altered by overexpressing the full length CcmM or CcmM-35 version resulted in an alteration in carboxysome size in *Syn*7942[114]. In these studies, a reduction in CcmM-58 relative to CcmM-35 levels resulted in larger carboxysomes, whereas increased CcmM-58 relative to Ccm-35 levels were correlated with smaller carboxysomes[114]. Notably, another band at approximately 30 kDa was detected by the anti-CcmM antibody (i.e., M* in Figure 2.4A) and accumulated to higher levels in the Δ*rcaE* mutant relative to WT. We conducted protein sequence analysis to determine the identity of this

Table 2.2: RNA sequencing data for carboxysome genes from *F. diplosiphon* SF33 WT and Δ*rcaE* mutant strains grown under GL or RL conditions.

Gene[a]	*Ava*[b] homo-log	No. Reads				Fold change[c] RL vs. GL		Fold change[d] Δ*rcaE* vs. WT	
		WT		Δ*rcaE*		WT	Δ*rcaE*	GL	RL
		GL	RL	GL	RL				
ccmK2	Ava_4472	1439	775	2954	2406	0.5*	0.8	2.1**	3.1**
ccmK1	Ava_4471	791	696	1612	2592	0.9	1.6*	2.0**	3.7**
ccmL	Ava_4470	362	391	776	1336	1.1	1.7*	2.1**	3.4**
ccmM	Ava_4469	1631	1668	2934	6722	1.0	2.3**	1.8*	4.0**
ccmN	Ava_4468	1061	1409	2073	3137	1.3	1.5	2.0**	2.2*
ccmO	Ava_4467	951	6310	639	6768	6.6**	10.6**	0.7**	1.1
lysR	Ava_4466	366	397	344	452	1.1	1.3	0.9	1.1
rbcL	Ava_3907	11568	13245	8840	29880	1.1	3.4**	0.8*	2.3
rbcX	Ava_3906	4522	5419	1249	2586	1.2	2.1**	0.3**	0.5
rbcS	Ava_3905	3828	6487	1390	3087	1.7	2.2**	0.4**	0.5
ccmK3	Ava_4709	451	371	396	471	0.8	1.2	0.9	1.3
ccmK4	Ava_4710	329	153	519	428	0.5**	0.8	1.6	2.8**
ccmP	Ava_4911	32	37	48	34	1.1	0.7	1.5	0.9
ccaA	Ava_2165	70	61	51	178	0.9	3.5**	0.7*	2.9**

a Colors indicate which genes are near each other in the genome (black are isolated from others in distinct regions of genome for each.
b ORFs were compared against *Anabaena variabilis* (*Ava*) ATCC 29413 annotated proteins using BlastX with a cut-off e-value of 0.0001 to determine *Ava* homolog.
c Fold change, differential expression analysis between two light treatments was calculated for each strain.
* $p < 0.5$ or ** $p < 0.01$ is significance value calculated for RL vs. GL counts for each strain.
d Fold change, differential expression analysis between two strains was calculated for each light condition.
* $p < 0.5$ or ** $p < 0.01$ is significance value calculated for Δ*rcaE* vs. WT counts for each light condition.

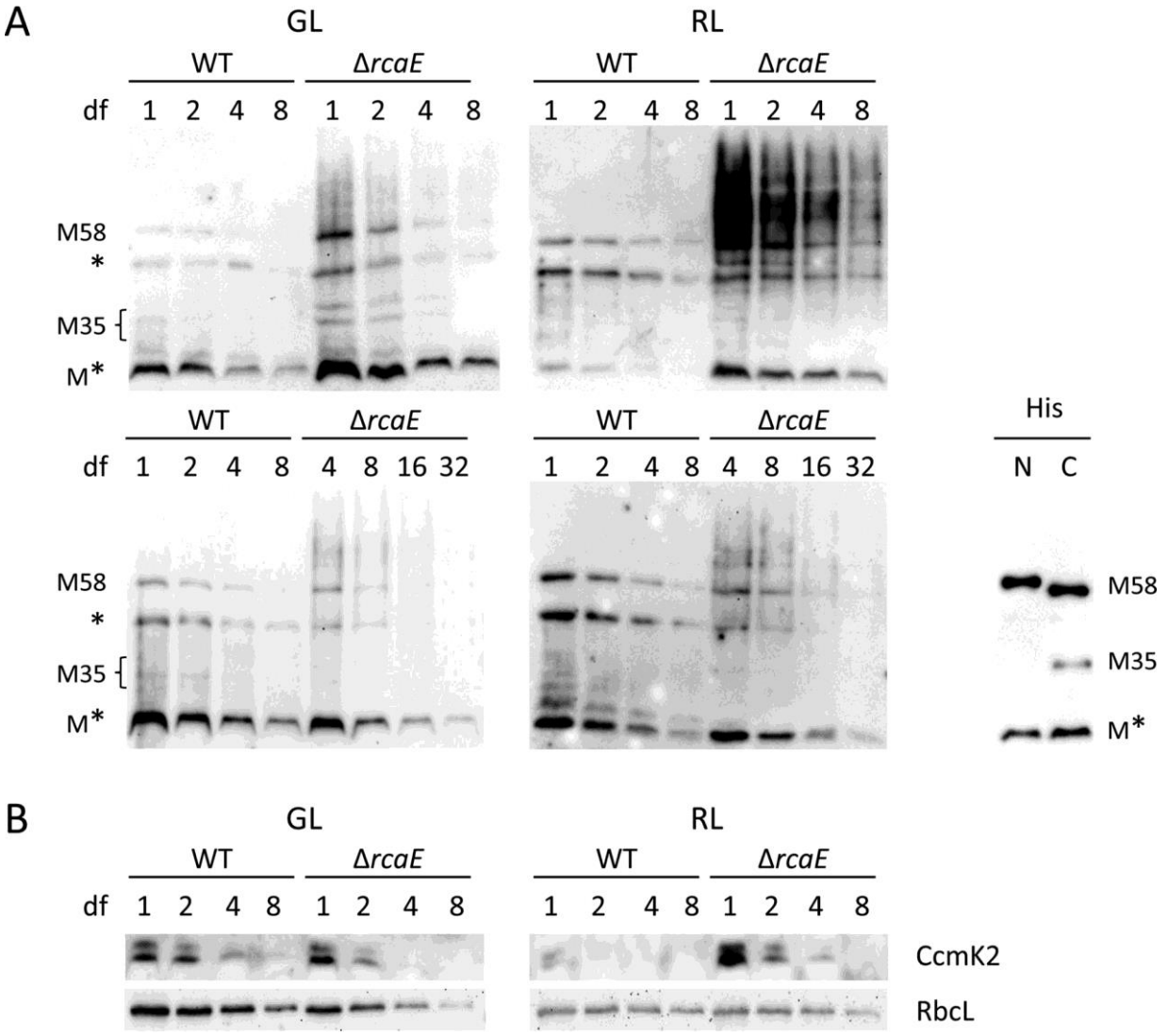


Figure 2.4: Immunoblot analyses of carboxysome protein accumulation in *F. diplosiphon*. Ccm protein accumulation in SF33 WT strain and Δ*rcaE* strains under GL and RL conditions are shown in representative blots. Protein extracts (μg of undiluted total protein extract indicated in parentheses) and two-fold dilutions as indicated by dilution factor (df, numbers above lanes) were loaded for assessment of **(A)** CcmM (75 μg) and **(B)** CcmK2 (75 μg) and RbcL (20 μg). After blotting, proteins were detected using anti-CcmM (1:5000 dilution, 3 min exposure), anti-CcmK2 (1:3000 dilution, 1 min exposure), or anti-RbcL (1:20000 dilution, 4 min exposure) antibodies. For panel A, distinct CcmM variants are indicated, which include *F. diplosiphon* versions of full-length CcmM-58 (M58), Ccm-35 (M35) that is derived from an internal ribosome entry site[114], and a reproducibly observed ~30 kDa band that we designated M* which is also observed in *Escherichia coli* expressed, N-terminal or C-terminal His-tagged versions of CcmM purified via Ni-NTA affinity chromatography (Panel A, lower right). M* protein was sequenced and determined to contain peptides which map to regions throughout the full-length Ccm protein (see Figure 2.5). *, indicated a non-specific band detected with the anti-CcmM antibodies. The lower portion of panel A with increased dilutions of soluble protein from the Δ*rcaE* strain was included to allow comparison of WT and Δ*rcaE* in a range of protein concentrations that was not saturating for Δ*rcaE*, given its significantly higher accumulation of CcmM-reactive bands in the same dilution range shown in upper portion of the panel.

band and determined that peptides for this protein map throughout the full length CcmM sequence (Figure 2.5), indicating overaccumulation of CcmM-derived bands in Δ*rcaE*.

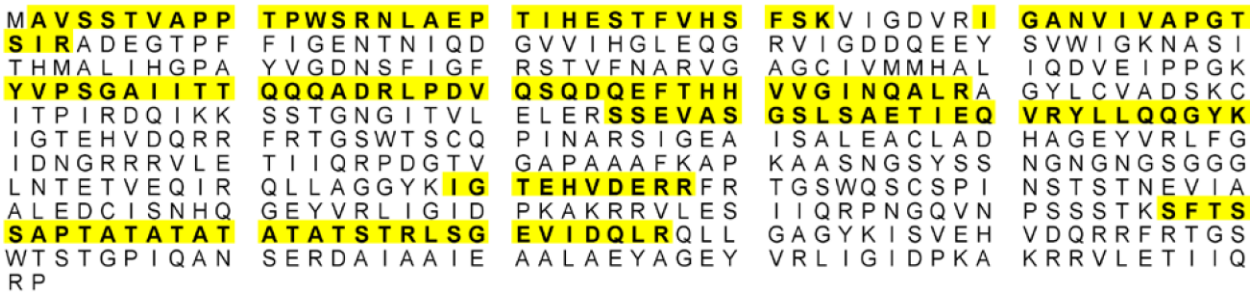


Figure 2.5: Protein sequencing of M* band. Highlighted regions represent exclusive peptides of the 30-kDa band identified in anti-CcmM immunoblots which map to the *F. diplosiphon* CcmM sequence. Nine exclusive unique peptides and 12 exclusive unique spectra were identified among 86 total spectra. The identified peptides map 152/552 amino acids (i.e., 28% coverage) of full-length CcmM.

As expected based on transcriptional downregulation of *rbcL* in GL, RbcL protein levels were significantly lower in the Δ*rcaE* mutant under these conditions (Figure 2.4B). Based on densitometry analysis, RbcL levels were reduced by 64% (SD = 0.07, n = 5) in *ΔrcaE* compared to WT in GL conditions. Under RL, RbcL protein levels may either decrease slightly or remain roughly constant, in contrast to the observed transcriptional upregulation. We observed a 31% reduction (SD = 0.31, n = 6) in RbcL levels in *ΔrcaE* compared to WT in RL. Together, these findings suggest an overaccumulation of carboxysome shell proteins and the CcmM protein which functions in nucleating the cargo relative to the levels of the carboxysome rubisco cargo, which are lower in Δ*rcaE* mutant compared to WT.

2.4.3 RcaE regulates the response of carboxysome structure to changes in light quality and intensity in *F. diplosiphon*

Due to the ability of photoreceptors such as RcaE to respond to light quantity in addition to light quality[15,161,164] and the prior correlation of increased carboxysome numbers under increased light

intensity in a cyanobacterium[199], we assessed carboxysome structure in WT and Δ*rcaE* strains under a variety of light conditions. Using TEM-based analyses, we measured carboxysome diameter and number/cell section for both strains grown under ~30 μmol m^{-2} s^{-1} medium RL (MRL) or GL (MGL). The Δ*rcaE* strain retained a small carboxysome phenotype relative to WT under both MRL and MGL (Table 2.1). However, the higher light intensity resulted in statistically larger carboxysomes in Δ*rcaE* under both MRL and MGL conditions, as well as in WT under MGL, when compared to standard (i.e., 10-15 μmol m^{-2} s^{-1}) light conditions (Table 2.1). However, no difference in size was noted when comparing WT grown under MGL to those grown under MRL or in WT under MRL compared to standard RL (Table 2.1). Thus, the loss of RcaE leads to light-dependent changes in all conditions, even if no changes are observed in WT.

2.4.4 RcaF and RcaC do not function with RcaE in the regulation of carboxysome-associated gene expression in *F. diplosiphon*

RcaF and RcaC function downstream of RcaE in the regulation of pigmentation[90]. Although, not required for RcaE-dependent regulation of morphology in GL, RcaF and RcaC contribute to morphology regulation under RL conditions[24]. To determine whether these effectors function downstream of RcaE in the regulation of carboxysome structure in cells, we assessed carboxysomes in Δ*rcaF* and Δ*rcaC* mutants. However, carboxysomes were very similar in appearance in these mutants to those in WT cells (Figure 2.6). Thus, RcaE-dependent regulation of carboxysomes does not occur through known response regulators RcaF or RcaC, as carboxysomes do not differ significantly from WT in either Δ*rcaF* or Δ*rcaC* mutants. In additional support of this TEM-based observation, the expression of *ccm* genes is not altered

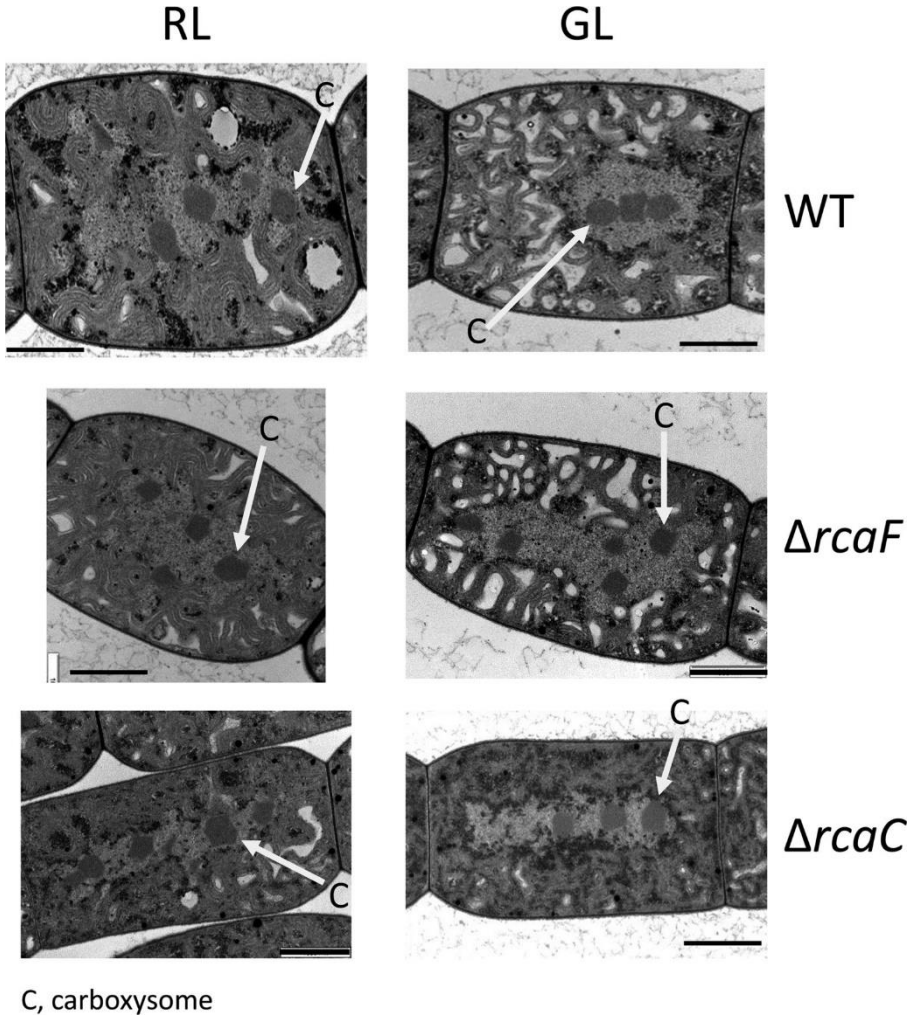


Figure 2.6: TEM analysis of cellular morphology of *F. diplosiphon* strains under GL and RL conditions. Representative images of WT pigmentation strain (top), Δ*rcaF* mutant strain (middle), and Δ*rcaC* mutant strain (bottom). C, carboxysomes indicated by white arrows. Bar represents 1 µm.

significantly in either Δ*rcaF* or Δ*rcaC* (Figure 2.7). Thus, RcaE appears to function primarily through unknown effectors to regulate carboxysome structure.

2.4.5 RcaE-dependent regulation of cell shape and intracellular reactive oxygen species levels are not correlated with the regulation of carboxysome structure in *F. diplosiphon*

Initial assays indicated that, in addition to being smaller and more numerous in Δ*rcaE* relative to WT, carboxysomes occasionally were mislocalized among thylakoid membranes rather than the expected location in the cytosol in Δ*rcaE* mutant cells (Figure 2.2A). Prior studies indicated

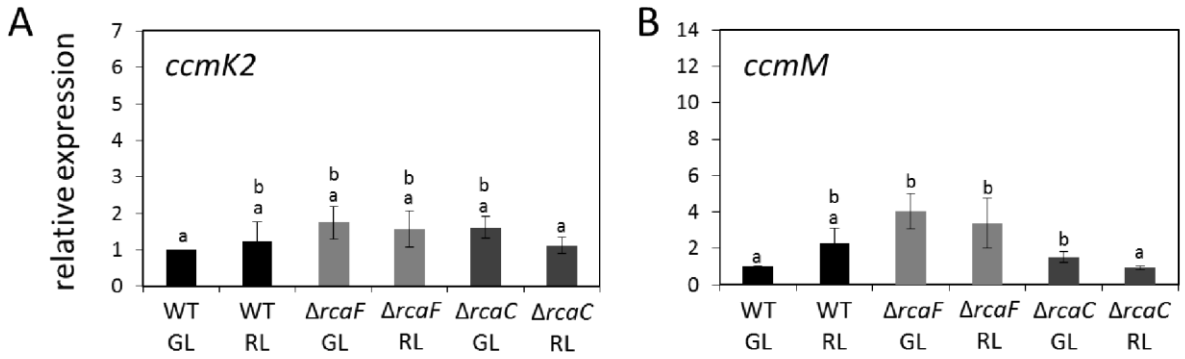


Figure 2.7: qPCR-based gene expression analyses in WT, Δ*rcaF*, and Δ*rcaC* strains of *F. diplosiphon*. Expression levels of **(A)** *ccmK2* and **(B)** *ccmM* genes in WT, Δ*rcaF*, and Δ*rcaC* strains grown under GL or RL are shown. Expression data for genes are presented relative to the internal control *orf10B*, and the data in each panel are shown relative to expression of the gene of interest in WT cells under GL conditions. Bars represent averages (± SD) of data from three independent biological replicates. Identical letters over bars represent homogenous mean groups ($p > 0.05$); different symbols indicate a statistically significant difference ($p < 0.05$) from others.

movement of carboxysomes from the central cytoplasm to the cell periphery under conditions of low inorganic carbon levels[127]. To determine whether this mislocalization phenotype or the observed carboxysome structural defect phenotype were primarily correlated with the spherical cell shape of the Δ*rcaE* mutant or other parameters, we assessed another spherical mutant of *F. diplosiphon*, i.e., Δ*bolA*[189]. BolA is a morphogene shown to be involved in regulation of cell shape in a number of bacteria[3,173]. The deletion of *bolA* is associated with large spherical cell shape in a Δ*bolA* mutant of *F. diplosiphon*[189] and its overexpression induces spherical cell morphology[190]. The Δ*bolA* mutant exhibited WT-sized carboxysomes; yet, these structures occasionally were mislocalized and found closer to the periphery of the thylakoid membranes, rather than centrally in the cytosol, in cells (Figure 2.8).

Given the prior recognition that redox state of the cell may impact CcmM activity[148], we investigated whether high ROS levels characteristic of Δ*rcaE* cells may contribute to the observed disruptions in CcmM levels and carboxysome phenotypes observed in this strain. As previously reported Δ*rcaE* cells accumulated elevated levels of ROS[186,188] (Figure 2.9). Thus, we

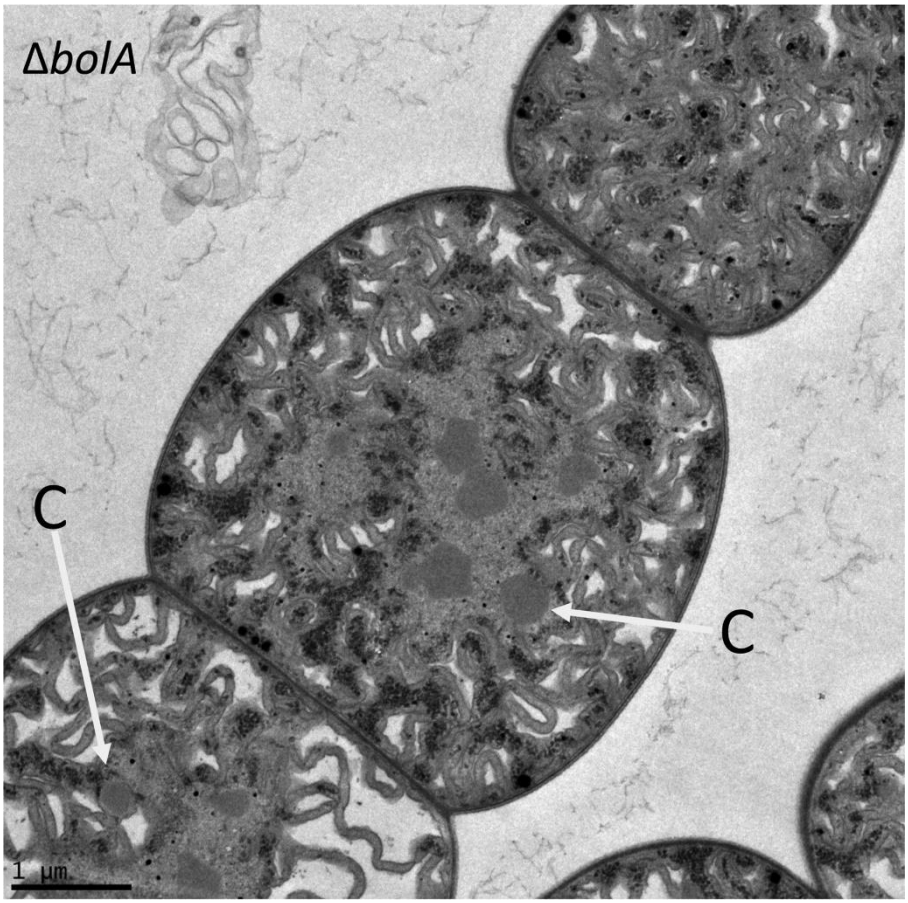


Figure 2.8: TEM analysis of ultrastructure of Δ*bolA* strain of *F. diplosiphon*. One representative image is shown. C, carboxysomes indicated by white arrows. Bar represents 1 µm.

investigated whether intracellular ROS accumulation is correlated with the smaller carboxysomes apparent in the Δ*rcaE* strain. To investigate the potential role of ROS in regulating carboxysome size, we treated Δ*rcaE* and WT cells with the ROS-mitigating antioxidant ascorbic acid (AA)[186]. AA-treated Δ*rcaE* cells exhibited reduced intracellular ROS levels compared to the untreated parental Δ*rcaE* strain[186] (Figure 2.9A). However, there were no significant differences between the size of carboxysomes in Δ*rcaE* in the presence or absence of AA, and carboxysomes were significantly smaller than WT controls in each case (Figure 2.9B & 2.9C). Furthermore, the aforementioned Δ*bolA* strain also has elevated intracellular ROS levels[189,190], which are not correlated with a change in carboxysome size in this strain relative to WT (Figure 2.8).

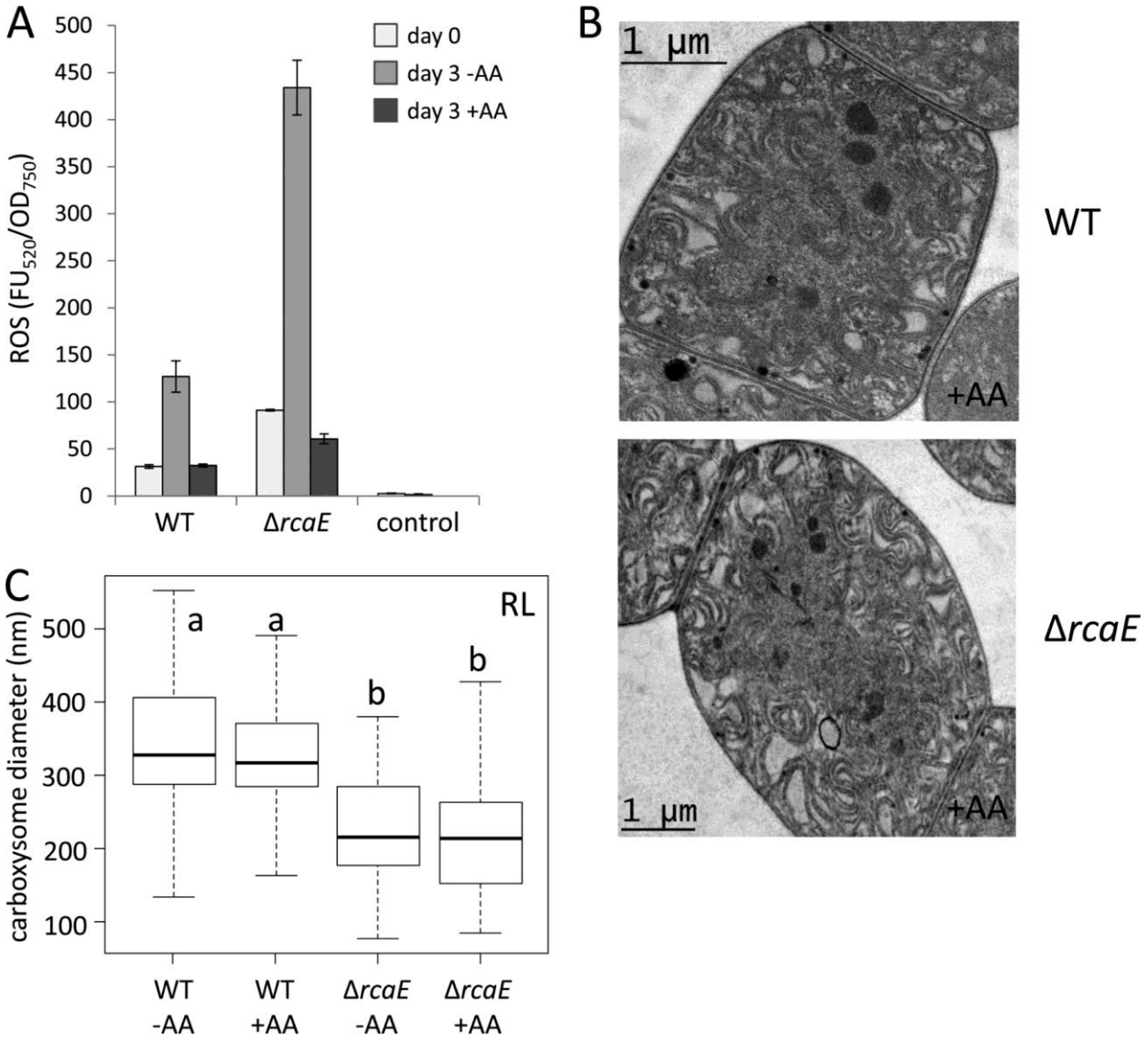


Figure 2.9: Accumulation of ROS and carboxysome structure in *F. diplosiphon*. (A) ROS-dependent DCF fluorescence and cell component fluorescence in *F. diplosiphon* under RL after 3 d treatment with (+AA) or without (-AA) ascorbic acid (2 mM) added to the growth medium in SF33 and Δ*rcaE* strains. The control represents BG-11 medium + DCFH-DA fluorescence without cells. Bars represent the average of FU, arbitrary fluorescent units at 520 nm relative to OD_{750} (FU_{520}/OD_{750}). Error bars indicate SD. **(B)** TEM analysis of strains under RL grown in presence of AA. Bars represent 1 µm. **(C)** Carboxysome size measurements of WT and Δ*rcaE* strains under RL, with and without AA. To determine size, the maximum diameters of at least 25 carboxysomes were measured from each strain under each growth condition and presented as a boxplot, with the bold line signifying the median diameter, the box representing the 2nd and 3rd quartile groups (the middle 50% of all values), and the lower vertical line corresponding to the 1st quartile group (the smallest 25% of all values) and the upper line corresponding to the 4th quartile group (the largest 25% of all values). Corresponding averages (± SE) can be found in Table 2.1. Statistical analyses were conducted using a Welch two sample *t* test performed in R. Identical letters over bars represent a homogenous mean group ($p > 0.05$), different symbols indicate a significant difference ($p < 0.05$) from others.

2.4.6 The structures of polyphosphate bodies are also regulated by RcaE in *F. diplosiphon*

To independently assess whether the smaller and greater number of carboxysomes per cell section of a Δ*rcaE* mutant observed in thin-section TEM analysis represent a smaller size, yet larger number of total carboxysomes in cells or an alteration in total carboxysome volume, we attempted to assess the whole cell population of carboxysomes. We used negative-staining of whole cyanobacterial cells with TEM analysis[97]. Results from these analyses indicated a larger number of smaller electron-dense bodies that appeared to have the potential shape of carboxysomes in Δ*rcaE* cells compared to WT, especially in RL (Figure 2.10). To confirm the identity of these structures, we used negative whole cell staining of *Syn*7942 WT cells and a carboxysome-deficient strain as potential controls (Figure 2.11). Upon observing similar electron-dense bodies in both of these lines, we conducted combined Scanning TEM (STEM) and energy-dispersive X-ray spectroscopy (EDX) elemental analysis methods to identify the smaller, more numerous structures apparent in Δ*rcaE* cells. EDX analyses indicated that the bodies observed were polyphosphate bodies (PPB) (Figure 2.12). This finding indicated an unexpected role for RcaE in regulating both carboxysome and PPB size and abundance.

In bacteria, *ppk* (encoding polyphosphate kinase 1) and *ppx* (encoding exopolyphosphatase) are involved in synthesis and degradation of polyphosphate (see review[29]). Notably, a Δ*ppk* mutant in *Synechococcus* lacked similar dark bodies in negative-staining TEM analysis to those we observed in *F. diplosiphon*[63]. Additionally, the Δ*ppk* had altered regulation of several *ccm* genes compared to WT, indicating a potential functional correlation between disruptions in PPB formation and carboxysome synthesis[63]. Given these observations, we assessed whether *ppk* and *ppx* mRNA levels were altered in our RNA-seq data. However, there were no significant

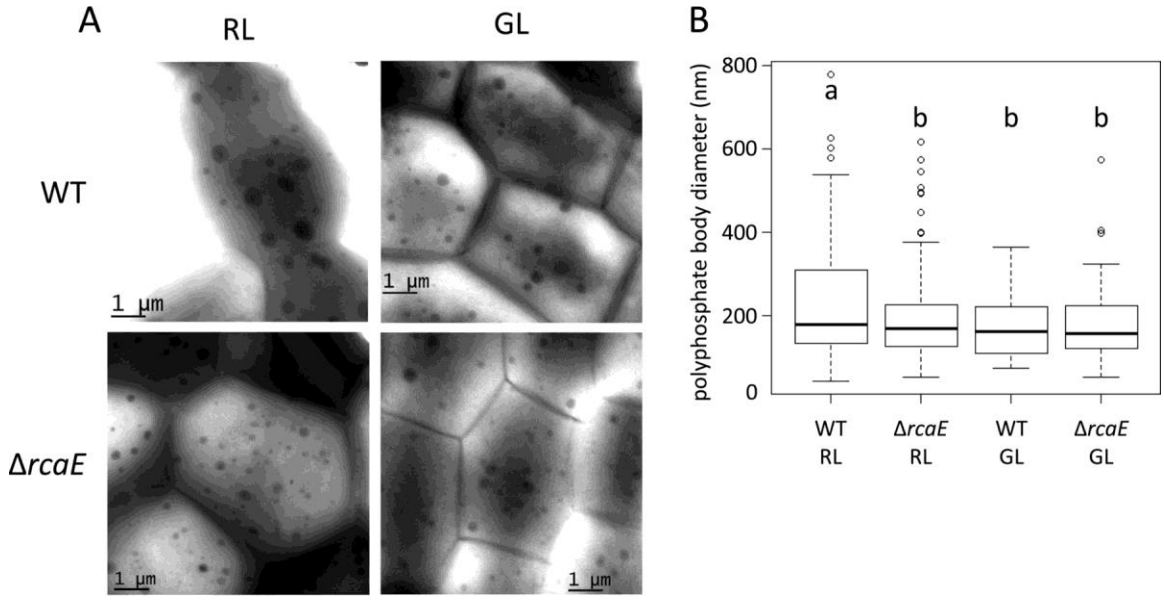


Figure 2.10: Assessment of PPB structure via TEM of whole cells of *F. diplosiphon* strains. (A) Representative TEM images of SF33 WT strain and Δ*rcaE* mutant strain grown under RL and GL. Dark spots are electron-dense bodies. **(B)** Quantification of diameter of PPB. To determine size, the diameters of at least 25 PPBs were measured from each strain under each growth condition and presented as a boxplot, with the bold line signifying the median diameter, the box representing the 2nd and 3rd quartile groups (the middle 50% of all values), and the lower vertical line corresponding to the 1st quartile group (the smallest 25% of all values) and the upper line corresponding to the 4th quartile group (the largest 25% of all values). Statistical analyses were conducted using a Welch two sample *t* test performed in R. Identical letters over bars represent a homogenous mean group ($p > 0.05$), different symbols indicate a significant difference ($p < 0.05$) from others.

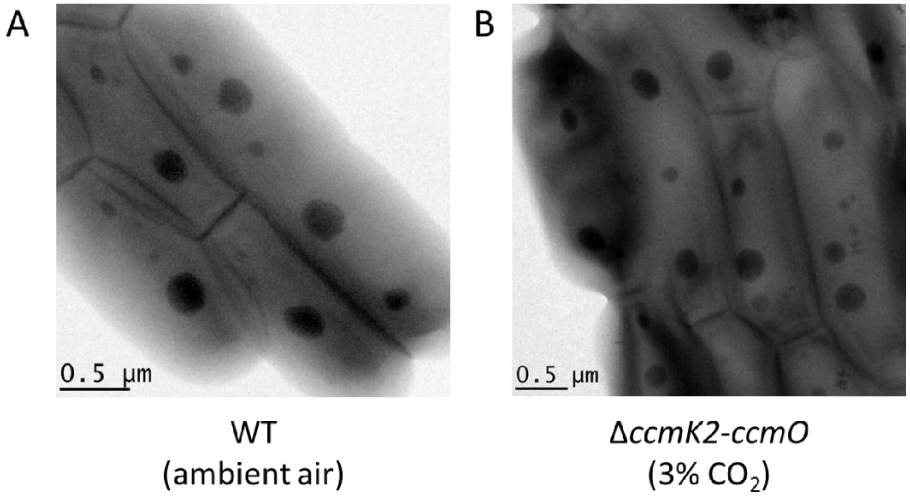


Figure 2.11: Negative staining via TEM of whole cells of a *Syn*7942 WT strain and a carboxysome-deficient Δ*ccmK2-ccmO* strain grown under WL conditions. To assess whether the electron-dense bodies observed in negative-stained whole cells were carboxysomes or other subcellular structures, we compared TEM images of **(A)** the WT strain, which was grown in ambient air, and **(B)** a carboxysome-deficient Δ*ccmK2-ccmO* strain, which has a high-carbon-requiring growth phenotype and thus was grown in 3% CO_2.

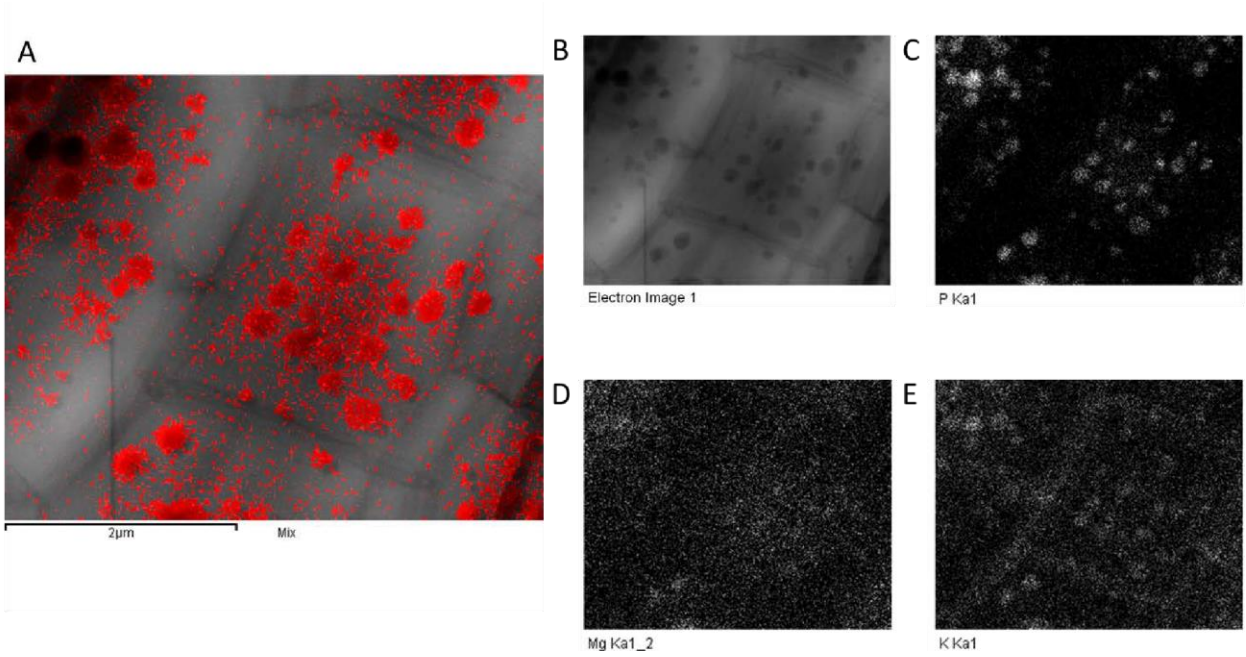


Figure 2.12: EDX spectroscopy elemental analysis in STEM mode of negative-stained whole-cell *F. diplosiphon* Δ*rcaE* mutant strain grown under RL. Multiple elements were analyzed for colocalization with electron-dense bodies. **(A** and **C)** Phosphate (P). **(B)** Electron-dense particles seen in STEM analysis. **(D)** Magnesium (Mg). **(E)** Potassium (K). The bar represents 2 µm and is applicable to panel A only.

differences in the mRNA levels of these genes accumulating in Δ*rcaE* vs. WT (Table 2.3). Thus, RcaE appears to regulate carboxysomes through transcriptional control of carboxysome genes; yet, the disruption in PPBs in Δ*rcaE* occurs without significant regulation of expression of genes known to impact PPB formation.

2.4.7 Total carboxysome population size or volume is regulated by RcaE in *F. diplosiphon*

As an alternative method to negative-stain TEM analysis to determine whether the smaller carboxysomes of Δ*rcaE* represent a smaller size, yet larger total population of carboxysomes in cells, we counted the number of carboxysomes observed in series of TEM thin sections and estimated total number per cell based on prior methods[35]. These analyses indicated that indeed the Δ*rcaE* mutant has a larger number of smaller carboxysomes per cell section than does WT under both GL and RL conditions (Figure 2.2C, Table 2.1).

Table 2.3: RNA sequencing data for polyphosphate synthesis and degradation genes from *F. diplosiphon* SF33 WT and Δ*rcaE* mutant strains grown under GL or RL conditions.

Gene	*Ava*[a] homo-log	No. Reads				Fold change[b] RL vs. GL		Fold change[c] Δ*rcaE* vs. WT	
		WT		Δ*rcaE*		WT	Δ*rcaE*	GL	RL
		GL	RL	GL	RL				
ppk	Ava_3165	474	349	559	536	0.7	1.0	1.2	1.5
ppx	Ava_3530	410	298	408	427	0.7	1.0	1.0	1.4

[a] ORFs were compared against *Ava* ATCC 29413 annotated proteins using BlastX with a cut-off e-value of 0.0001 to determine *Ava* homolog.
[b] Fold change, differential expression analysis between two light treatments was calculated for each strain. Note: no significant differences were detected using *t*-test.
[c] Fold change, differential expression analysis between two strains was calculated for each light condition. Note: no significant differences were detected using *t*-test.

Given the correlation between changes in light intensity with changes in carboxysome size, we also assessed whether the number of carboxysomes increased under increased light intensity. Light intensity typically did not alter the average carboxysome number per cell in either strain, except for a slight decrease in WT under MGL (Table 2.1). Similarly, there were no significant differences in carboxysome number when comparing standard RL and GL in either strain. This suggests that the number of carboxysomes per cell is well-maintained in WT, that the number is dependent on the presence of RcaE in *F. diplosiphon*, and that carboxysome size is primarily sensitive to dynamic photoenvironments.

2.5 Discussion

Here, we report a regulatory role for RcaE in maintaining carboxysome size, quantity per cell, and contributing to carboxysome subcellular localization in *F. diplosiphon*. The Δ*rcaE* mutant, which lacks cyanobacteriochrome RcaE photoreceptor[89], has smaller and more numerous

carboxysomes than the parental WT line. These observations provide evidence that RcaE contributes to the regulation of carboxysome size and quantity in *F. diplosiphon*. Carboxysomes also are mislocalized occasionally within the thylakoid membranes of this strain. Notably, prior studies have reported a shift in location of carboxysomes from the central cytoplasm to the periphery of the cell under levels of reduced C_i availability[127].

PPB morphology also is disrupted in the Δ*rcaE* strain, with more numerous and smaller PPBs than observed for WT cells. This is notable given several prior recognized associations between phosphate-rich PPBs and carboxysomes in bacteria. In one study in a proteobacterium, the position of PPBs correlated with the positioning and structure of carboxysomes[80]. More closely related to the work here, carboxysomes have been previously reported to be closely associated or grouped with PPBs in some cyanobacterial strains[109,138]. Whether these associations point to functional interaction remains to be definitively determined; however, misregulation of *ccm* genes in a *Synechococcus* mutant lacking PPBs hints at a functional association[63]. Given the noted association of PPBs with DNA in the cytoplasm[134,140,181], the physical co-localization may indicate that carboxysomes are also nearby or associated with DNA. This association allows for a number of possible connections between chromosome condensation/de-condensation dynamics, gene expression regulation, and subcellular structures to be explored, especially since the Δ*rcaE* mutant exhibits disruptions to both carboxysome and PPB morphology. Our observations that only some *ccm* genes are significantly misregulated, that *ccm* genes in disparate regions of the genome are misregulated, and that there is no apparent change in expression of *ppk* and *ppx* genes which are associated with PPB synthesis in the Δ*rcaE* mutant, suggest that disruptions to carboxysome and PPB structures do not arise from non-specific changes to

chromosome accessibility in the nucleoplasm. However, the shared structural phenotypes of carboxysomes and PPBs, alongside their previously reported associations, most likely highlight a robust interconnectivity (perhaps functional) between these subcellular structures. Taken together, we hypothesize that RcaE could have a role in multiple aspects of carboxysome regulation, including interactions with PPBs, that are likely critical for carboxysome dynamics and function in carbon fixation in a cyanobacterium.

RcaE was previously described as the photosensory receptor that controls pigmentation and cellular morphology in *F. diplosiphon*[23,89,209]. RcaE works through two known response regulators RcaF and DNA-binding transcriptional regulator RcaC in regulating pigmentation[4,21,40,89,90,104–106] and red-light-dependent regulation of cellular morphology[24]. Notably, however, RcaE does not appear to function through RcaF and RcaC in the regulation of carboxysomes, as Δ*rcaF* and Δ*rcaC* mutants have no apparent defects in the regulation of carboxysome size or positioning, nor significant misregulation of expression of major *ccm* genes, *ccmM* or *ccmK2*. In the regulation of cellular morphology, RcaE controls expression of the *bolA* morphogene[189,190]. Yet, RcaE also does not impact carboxysome morphology via BolA regulation as a Δ*bolA* mutant has WT-like carboxysomes. Thus, although RcaE impacts expression of carboxysome genes and carboxysome structure and number, the effectors through which it functions to do so appear independent of known RcaE-regulated effectors controlling pigmentation and cell shape phenotypes characteristic of CCA.

Of note, localization of carboxysomes generally may be correlated with cell shape, as carboxysomes are mislocalized to the periphery of cells in mutants with a constitutive spherical

morphology, including both the Δ*bolA* mutant and the Δ*rcaE* strain. Previously, there have been additional correlations made between carboxysomes and cell shape. Elongated cell division mutants exhibit decreased carboxysome numbers per cell and carboxysome structural defects[65]. Notably, these mutants also have reduced levels of carboxysome-associated proteins[97]. Additionally, impairments in cell morphology due to cytoskeleton defects were correlated with altered spatial distribution or mislocalization of carboxysomes in cells[174].

In addition to its spherical morphology, the Δ*rcaE* mutant has elevated intracellular ROS levels[23,186]. Despite both Δ*rcaE* and Δ*bolA* strains having elevated ROS[186,189], Δ*rcaE* mutant cells have smaller carboxysomes and Δ*bolA* mutant cells have WT-sized carboxysomes. Additionally, even when intracellular ROS levels were reduced in Δ*rcaE* mutant cells treated with an antioxidant, carboxysomes were smaller in cells lacking RcaE. Thus, RcaE appears to have a direct regulatory role in controlling carboxysome morphology and dynamics, rather than indirectly impacting carboxysomes through altering intracellular ROS accumulation.

The regulatory role of RcaE related to carboxysome structure and function is linked to transcriptional regulation of *ccm* and carboxysome-associated genes. In a Δ*rcaE* mutant, Ccm structural proteins overaccumulate and carboxysome cargo protein, rubisco, underaccumulates relative to levels in WT. This observed shift in the carboxysome protein profile results in reduced cargo and simultaneous elevated levels of the shell and CcmM-58, which may contribute to the generation of smaller, more numerous carboxysomes.

F. diplosiphon is able to adjust carboxysome size to a number of changes in its photoenvironment. Carboxysomes in the WT strain respond to increased light availability through an increase in size, with this effect being more pronounced under GL. Carboxysomes also appear to be larger under red light compared to GL at low light levels, but this effect is lost at higher light intensities. These data are consistent with a higher linear electron flow driving a larger need for carbon fixation. Since these general behaviors are not entirely lost in the Δ*rcaE* strain, more cellular factors are implicated in the light-dependent regulation of carboxysome structure. However, the loss of RcaE severely limits the maximum size of carboxysomes while increasing their number in all light conditions studied. Moreover, more light-dependent differences in carboxysomes are observed in the Δ*rcaE* strain where WT shows limited light-dependence, suggesting that RcaE is required to maintain carboxysome homeostasis under dynamic photoenvironments. This tendency to regulate carboxysome structure encourages future analyses of the extent to which RcaE-dependent alterations to carboxysome size and distribution can specifically impact carbon fixation.

Together, these results suggest that RcaE has a critical role in regulating carboxysome structure, which likely serves to match carbon fixation potential with external light cues. Given the light-dependent regulation of expression of *ccm* genes in RL vs. GL and altered levels of expression of *ccm* and carboxysome-associated genes in the RcaE-deficient strain relative to WT, both the structure (i.e., size and quantity) and composition (e.g., elevated *ccmL*, *ccmM*, *ccmO* and *ccmN* in RL vs. GL, and reduced levels of RbcL in GL vs. RL) of carboxysomes appear to be regulated and, indeed, fine-tuned in response to external light cues. Such a role for RcaE provides a key mechanism for matching carbon-fixation capacity and photosynthetic potential of cells to

available light. Given the prior observation that light intensity also regulates carboxysome structure and dynamics[76,79,199] and that phytochrome-related photoreceptors respond to light intensity in addition to light quality[168,169,211], we propose that RcaE plays a central role in tuning the structure and function of carboxysomes in response to a dynamic photoenvironment to optimize organismal fitness in *F. diplosiphon*.

2.6 Materials and methods

2.6.1 Culture conditions

Two strains of *F. diplosiphon* were compared in this study: a short-filament wild-type pigment strain (hereafter WT), which was identified as SF33[42], and a RcaE-deficient mutant strain (Δ*rcaE*)[89]. Strains were grown in BG-11 medium (Fluka, Buchs, Switzerland) with 20 mM HEPES (pH 8.0) (hereafter referred to as BG-11/HEPES) at 28°C with continuous shaking at 175 rpm under continuous light conditions. Liquid starter cultures were inoculated from strains maintained on solidified BG-11/HEPES media (BG-11/HEPES containing 1.5% [w/v] agar) and grown under continuous white fluorescent light (General Electric; model no. F20T12/PL/AQ/WS) at ~15 µmol m^{-2} s^{-1}. Exponentially growing cultures were diluted to an initial optical density at 750 nm (OD_{750}) of ~0.05 and were transferred to the experimental culture condition as indicated. Absorbance measurements were made with a SpectraMax M2 spectrophotometer (Molecular Devices, Sunnyvale, CA).

RL and GL conditions were obtained using monochromatic growth chambers at an intensity of ~10 to 12 µmol m^{-2} s^{-1} continuous broad-band RL (CVG sleeved Rosco red 24 fluorescent tubes, General Electric; model no. F20T12/R24) or continuous broad-band GL (CVG sleeved Rosco

green 89 fluorescent tubes, General Electric; model no. 20T12/G78) as previously described[23]. Growth under medium light intensity utilized ~30 µmol m^{-2} s^{-1} continuous RL (λmax 660 nm; model no. 2506RD LED Grow Light, LED wholesalers, Hayward, CA) or GL (λmax 530 nm; Sunbow model no. SN 1320001-004, Geneva Scientific LLC). For cultures grown to test supplemental carbon dioxide conditions, we used white light (WL) conditions in a Percival I-41LL incubator equipped with Phillips Alto II fluorescent lights (Model no. F17T81TL841) under either air or 3% CO_2. Light intensities were measured with a LI-250 Light meter (LI-COR, Lincoln, NE) equipped with a quantum sensor (model LI-190SA).

2.6.2 TEM and EDX analysis

2.6.2.1 TEM analysis of sectioned cells

For conventional TEM, ~10 mL of cells were harvested from exponentially growing cultures (i.e., OD_{750} at 0.6 – 0.8) by centrifugation at 5125 x *g* at room temperature for 6 min. Spent media was decanted, pellets were resuspended in remaining media (~ 200 µL), and transferred to 1.5 mL microfuge tubes. Cells were then centrifuged at 16,000 x *g* at room temperature for 5 min and remaining medium was removed. Cells were prefixed via resuspension with 1 mL of 2.5% (w/v) glutaraldehyde and 2.5% (w/v) paraformaldehyde in 0.1 M cacodylate buffer and incubated for 5 min at 33°C at 35% power in a Precision Pulsed Laboratory Microwave 9000 (Electron Microscopy Sciences, Hatfield, PA). After 3 x 10 min washes with 1 mL of 0.1 M cacodylate buffer, the cell pellets were resuspended in 2% (w/v) molten agarose in dH_2O and then centrifuged for 1 min. The solidified agarose plug was removed from the microfuge tube and the dense, embedded cell pellet was cut into ~1-4 mm cubes. Embedded cells were washed 3 x 10 min with cacodylate buffer and then postfixed with 2% (w/v) osmium tetroxide in

cacodylate buffer for 5 min at microwaved at 33°C, 35% power. Cells were then washed 3 x 10 min using 0.1 M cacodylate buffer followed by 3 x 10 min washes using dH_2O. Postfixed cells were blocked with 2% (w/v) uranyl acetate (Electron Microscopy Sciences, Hatfield, PA) in dH_2O and microwaved for 5 min at 33°C, 35% power, which has been reported to enhance contrast of carboxysomes[64]. Following 3 x 15 min washes with dH_2O, fixed cells were dehydrated in an acetone series (30%, 50%, 70%, 80%, 90%, 95%, 100%, 100%, 100%) in 20 min intervals using either an EMP5160 tissue processor (Boeckeler Instruments, Inc., Tucson, AZ) or manually. Dehydrated samples were infiltrated with Spurr resin (Firm Standard; Electron Microscopy Sciences, Hatfield, PA) in a 3:1, 2:2, 1:3 series of acetone:Spurr for 2-3 h at room temperature or overnight at 4°C at each step. Infiltrated cells were soaked in Spurr resin for 48 h with 3 exchanges of resin, then blocks were cured at 60°C for 48 h. Thin sections were prepared using a PowerTome XL ultramicrotome (Boeckeler Instruments, Inc., Tucson, AZ), and 70-90 nm sections (estimated from silver to gold interference color) were placed on 200 mesh Cu grids (Electron Microscopy Sciences, Hatfield, PA). Grids were stained with 4% (w/v) osmium tetroxide in dH_2O for 30 min followed by Reynold's formula (lead citrate, comprised of lead nitrate and sodium acetate; Electron Microscopy Sciences, Hatfield, PA) for 15 min while covered alongside NaOH pellets. Sections were imaged using a JEOL 100CX TEM (JEOL USA Inc., Peabody, MA) equipped with a MegaViewIII digital camera at an operating voltage of 100 V.

2.6.2.2 Negative-staining TEM analysis of whole cells.

For negative-staining TEM analysis of whole cells, 2 mL of cells at an $OD_{750} \geq 0.2$ or 5 mL of cells at an $OD_{750} \leq 0.2$ were harvested by centrifugation at 5125 x *g* at room temperature for 6

min. Spent media was decanted, pellets were resuspended in 15 mL of dH_2O, and then cells were centrifuged at 5125 x *g* at room temperature for 6 min. The supernatant was discarded by aspiration and the pellets were resuspended in 0.5 mL dH_2O. A 5-µL aliquot of resuspended cell pellet was placed on a 200 mesh Cu grid coated with Formvar (Electron Microscopy Sciences, Hatfield, PA) and incubated for 2 min at room temperature. The grid was blotted nearly dry with Whatman filter paper. Either 5 µL of 0.1% (w/v) uranyl acetate in dH_2O (stained) or 5 µL of dH_2O (unstained) was added to the grid and blotted away after 5 sec. Grids were then washed once with 5 µL of dH_2O for 5 sec, then blotted nearly dry. Grids were imaged using a JEOL 100CX TEM equipped with a MegaViewIII digital camera at an operating voltage of 100V.

2.6.2.3 EDX STEM analysis

EDX analysis was performed both on conventional TEM sections and negative-stained samples using a JEM-2200FS TEM (JEOL USA Inc., Peabody, MA) with an in-column energy filter operated in STEM mode at 200 kV. The analytical work was done with the attached Oxford Instrument INCA system with energy resolution of 140 eV. The images were collected with a Gatan Multiscan camera at 1024x1024 resolution.

2.6.3 Carboxysome and PPB size and number quantification

To determine size of carboxysomes and polyphosphate bodies (PPBs), the diameter of at least 25 of each were measured from each strain under each growth condition in TEM images. Analysis was done in the image editing software Paint.net, and we selected the maximum diameter for consistency between irregular shapes. The number of carboxysomes and PPBs were determined by counting positively identified structures in ~30 cell sections using TEM sections (for

carboxysomes) or ~ 10 cells using negative-staining TEM (for PPBs) for each strain in each condition. Positive identification of a carboxysome structure (for both size and number) satisfied three criteria: 1) appearance of some sharp edges, 2) moderate electron density in contrast to the cytosol, 3) regular, paracrystalline distribution of electron density within the carboxysome. The negative-staining technique highlighted cell outlines and allowed visualization of the naturally electron-dense PPBs. For both carboxysome and polyphosphate quantification, we used boxplots to display data. Boxplots were used as they present the entire data population spread, ordered from smallest to largest. The horizontal bold line inside each boxplot graph corresponds to the median, and the box covers the 2nd and 3rd quartile groups (the middle 50% of all values). The vertical line below the box corresponds to the 1st quartile group (the smallest 25% of all values) and the line above the box corresponds to the 4th quartile group (the largest 25% of all values). Presenting the entire spread of data allows for visualization of differences between population spreads. Averages (± SE) are also presented for carboxysome size and number in tabulated format.

2.6.4 qPCR analyses

The abundance of *ccmK1*, *ccmK2*, *ccmL*, *ccmM*, *ccmN*, *ccmO*, *ccmK3*, and *ccmK4* transcripts in total RNA extracted from GL- and RL-grown WT or Δ*rcaE* strains of *F. diplosiphon* were analyzed using the delta delta C_q ($\Delta\Delta C_q$) method as detailed previously[189]. In brief, total RNA was extracted as described[180,188] and reverse transcribed (0.5 μg in a 20 μL reaction mixture) with random primers using the Promega reverse transcription kit according to the manufacturer's instructions. No reverse transcription control reactions, which lacked reverse transcriptase enzyme in the reverse transcription reaction mixture, were also performed for all samples. After

reverse transcription, the reaction mixture was diluted with 30 µL of nuclease-free water and 3 µL of this reaction was used in a 10 µL total qPCR reaction according to manufacturer's instructions using the Microamp® fast optical 96-well reaction plate with barcode and ABI FAST 7500 Real-Time PCR System (Applied Biosystems) in FAST mode with Fast SYBR Green Master Mix (Applied Biosystems). Primers sets used for each gene and *orf10B* internal control, the latter expressed equally under GL and RL conditions[197], are listed in Table 2.4. The annealing/extension temperature for all primer sets was 60°C, and all primers were verified to produce a single product by melting curve analysis. The abundance of transcripts was determined based on relative quantification with normalization to the reference transcript *orf10B.* All qPCR experiments were performed with three independent biological replicates and three technical replicates for each biological replicate. All qPCR procedures and analyses were performed according to the MIQE guidelines[32].

Table 2.4: qPCR primers used in chapter 2.

Primer name	Forward primer/FP (5'-3')	Reverse primer/RP (5'-3')
ccmK1	AACGAATTGGCAGGACATACT	GCAGGCGTAGAATCTGTGAA
ccmK2	AGGCTTGCACTTCCGATAC	TGCTGATGCGATGGTGAA
ccmL	GTCTACTCCTGCACCTACGATA	GTCTTCGAGGTGTGAAACTACTG
ccmM	GATTGCTCCCGAAGGTACATATT	GGCTTTCGCTCTACGGTATTT
ccmN	TGGCACTCAGATTTATGGTACAG	GTCCGAGATGGGTTCATTTAGAG
ccmO	CCATTACCTCCAAGCTCAGTAAA	CTCCTACCATCGCTGGAAATC
ccmK3	TGCTGCTGGAGAACAAGTAAA	GTAAAGTGGATCGGAAGGATGG
ccmK4	CAGGCAGTTGGAGCATTAGA	TCAGAAACATCGCCACGAATA
orf10B	AGAACTACAGCGTCAGCTTAAT	CTGCTTCGCTTTCAGCATTT

2.6.5 Measurement of ROS

Reactive oxygen species (ROS) and other peroxide levels were measured using the fluorescent dye 2′,7′-dichlorodihydrofluorescein diacetate (DCF-DA; EMD chemicals, Gibbstown, NJ) according to previously described methods[74,186]. In brief, aliquots of cells were collected immediately after

dilution to a starting OD_{750} of 0.05 (Day 0) and after 72 h (Day 3) under desired growth conditions. In a dark room, aliquots were incubated with 10 µM (final concentration) of DCF-DA for 1 h at room temperature with rocking. Fluorescence measurements were then taken at 520 nm with excitation at 485 nm, using water to blank. The measurements were normalized by the OD_{750} of the culture and are directly proportional to total hydroxyl groups in the sample.

2.6.6 Protein extraction

After 7 d of growth in the desired condition, cells were harvested by centrifugation at 5125 x *g* at 4°C for 10 min. Spent media was decanted, then pellets were resuspended in remaining media (~ 200 µL) and transferred to 1.5 mL microfuge tubes. Cells were then centrifuged at 16,000 x *g* at 4°C for 5 min, the remaining media was aspirated, and the cell pellet mass was recorded using a Mettler Toledo XS104 Analytical Balance (Mettler Toledo, Columbus, OH). The pellets were resuspended in 20 mM Tris-HCl (pH 7.5) with 0.6 M sucrose[114], 0.2 mg/mL (w/v) lysozyme, 1X Protease Arrest (G Biosciences, St. Louis, MO), and 5 mM EDTA, at a ratio of 6 mL buffer per gram cell paste and transferred to 15 mL Falcon tubes. Samples were passed through a pre-chilled French Pressure Cell Press (SLM Instruments, Urbana, IL) at 500 PSI a total of three times per sample. Each sample was collected in a 15 mL Falcon tube and then transferred into 1.5 mL microfuge tubes and centrifuged at 16,000 x *g* at 4°C for 5 min. Following collection of the soluble fraction, the cell pellet was resuspended to the original volume using 20 mM Tris buffer to obtain a resuspended insoluble fraction of nearly equal concentration as the obtained soluble fraction.

2.6.7 Quantitative western blot analysis

Prior to SDS-PAGE, total protein concentration of soluble lysates was measured using the bicinchoninic acid (BCA) assay (Pierce™ BCA Protein Assay Kit; Thermo Fisher Scientific, Waltham, MA) following the manufacturer's recommendations. Lysates were then normalized by total protein, with addition of 20 mM Tris (pH 7.5) containing 0.6 M sucrose where needed. Samples normalized by total protein were then diluted with 5X SDS sample buffer, and then a 2-fold dilution series, up to a 32-fold dilution, was conducted using 1X SDS sample buffer. Insoluble fractions and whole cell pellets were resuspended in 1X SDS before loading. Samples were denatured at 95°C for either 1 min (soluble fractions) or 5-10 min (insoluble fractions).

Proteins (with expected kilodalton values for monomers shown in brackets) were separated on Tris-HCl gels with 10% acrylamide (CcmM [35 or 60 kDa for the short and long isoforms, respectively] or RbcL [53 kDa]) or 15% acrylamide (CcmK2 [11 kDa]) using Tris-Glycine SDS Running Buffer. After separation by electrophoresis, proteins were transferred to Immobilon-P polyvinylidene difloride membrane (EMD Millipore, Billerica, MA) using a semidry TransBlot Turbo Transfer System (Bio-Rad, Hercules, CA) at 25V (1.0 A max) for 40 min. PVDF membranes were blocked for 1 h at room temperature using 5% (w/v) dry milk in Tris-buffered saline with 0.5% (v/v) Tween-20. After blocking, membranes were probed using polyclonal rabbit antiserum raised against *Syn*7942 CcmK2[33] or CcmM[64] (α-Ccm antibodies provided by Dr. Cheryl Kerfeld), as well as antiserum raised against *Spinacia oleracea* RbcL (AS07 218, Lot 1004, Agrisera, Vännäs, Sweden). Primary antibody incubation was performed up to overnight at 4°C. Blots were washed 4 x 10 min in Tris-buffered saline with 0.1% (v/v) Tween before addition of goat-anti-rabbit, HRP-conjugated secondary antibody at a dilution of 1:20,000 for 1 h

at room temperature. Following four 10 min washes in Tris-buffered saline with 0.1% (v/v) Tween and two 5 min washes in Tris-buffered saline, HRP signal was detected using FEMTOGLOW™ Western PLUS HRP substrate (Michigan Diagnostics, Royal Oak, MI) on a ChemiDoc XP (Bio-Rad, Hercules, CA) imaging system.

2.6.8 Densitometry analysis

Densitometry was performed using ImageLab (Bio-Rad, Hercules, CA) software. Lanes were manually selected and bands were detected using high sensitivity, discarding bands that were clearly staining artifacts. Disc size, which determines the baseline, was set such that it reliably connected the bases of non-overlapping peaks (typically this was a disc size of 10 - 20). Using the same method, total protein was analyzed using Coomassie-stained gels run in parallel to western blots. The ratio of total protein in WT relative to Δ*rcaE* was analyzed for each dilution factor and found to be nearly 1 in the linear range.

2.6.9 Purification of *Fd*CcmM after expression in *E. coli*

Primers for *ccmM* from *F. diplosiphon* were designed with overhanging restriction sites such that the PCR fragment could be introduced into pET28a to add either a N-terminal (using restriction sites for *Nhe*I and *Xho*I) or C-terminal (using restriction sites for *Nco*I and *Xho*I) 6x His tag. After standard cloning methods and bacterial transformation, *E.coli* BL21strains containing each of the two constructs were analyzed to confirm fragment insertion. The expression of *ccmM* was induced overnight at 30°C using 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Cell pellets were harvested, resuspended in 15mL of native binding buffer (50 mM NaH_2PO_4, 0.5 M NaCl, pH 8.0) and passed two times through a CF Range Cell Disruptor (Constant Systems Ltd.,

Daventry, Northants, UK) operated at 15 kPSI in a 4°C cold room. Lysate was then spun at 5125 x *g* for 15 min at 4°C and the soluble fraction was extracted and incubated with Ni-NTA for 1 h in a purification column (Invitrogen Life Technologies, Carlsbad, CA). Affinity chromatography was performed according to manufacturer's instructions and bound protein was eluted using native binding buffer containing 250 mM Imidazole. SDS-PAGE analysis was used to identify elution fractions containing purified CcmM.

2.6.10 Protein sequencing analysis of ~30 kDa band identified from anti-CcmM immunoblot of *Fd*CcmM.

After electrophoresis of an elution fraction containing purified *F. diplosiphon* CcmM protein expressed in *E. coli* with an N-terminal His tag, gel bands ranging from 28-35 kDa were digested as previously described[182], with modifications, at the Proteomics Research Technology Support Facility at Michigan State University. Briefly, gel bands were dehydrated using 100% acetonitrile and incubated with 10 mM dithiothreitol in 100 mM ammonium bicarbonate (pH ~8) at 56°C for 45 min, dehydrated again and incubated in the dark with 50 mM iodoacetamide in 100 mM ammonium bicarbonate for 20 min. Gel bands were then washed with ammonium bicarbonate and dehydrated again. Sequencing grade, modified trypsin was prepared to 0.01 µg/µL (w/v) in 50 mM ammonium bicarbonate and ~50 µL of this was added to each gel band so that the gel was completely submerged. Bands were then incubated at 37°C overnight. Peptides were extracted from the gel by water bath sonication in a solution of 60%ACN/1%TCA (v/v) and vacuum dried to ~2 µL. Peptides were then re-suspended in 2% acetonitrile/0.1%TFA to 20 µL. From this, 5 µL were automatically injected by a Thermo EASYnLC 1000 (Thermo Fisher Scientific, Waltham, MA) onto a Thermo Acclaim PepMap RSLC 0.075 mm x 250 mm C18

column and eluted over 16 min with a gradient of 5% Buffer B (i.e., B; Buffer B = 99.9% Acetonitrile/0.1% Formic Acid) to 30% B in 1 min, ramping to 90% B at 2 min and held at 90% B for the duration of the run at a constant flow rate of 300 nL/min. Buffer A = 99.9% Water/0.1% Formic Acid. Eluted peptides were sprayed into a ThermoFisher Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA) using a FlexSpray spray ion source. Survey scans 2 were taken in the Orbi trap (35000 resolution, determined at m/z 200) and the top ten ions in each survey scan are then subjected to automatic higher energy collision induced dissociation (HCD) with fragment spectra acquired at 17,500 resolution. The resulting MS/MS spectra are converted to peak lists using Mascot Distiller, v2.6 (Matrix Scientific, Boston, MA) and searched against a database containing all cyanobacteria protein sequences and all *E.coli* protein sequences available from NCBInr (downloaded 2017-07-07 from www.ncbi.nlm.nih.gov) appended with common laboratory contaminants (downloaded from www.thegpm.org, cRAP project) using the Mascot searching algorithm, v 2.5. The Mascot output was then analyzed using Scaffold, v4.8.2 (www.proteomesoftware.com) to probabilistically validate protein identifications. Assignments validated using the Scaffold 1% FDR confidence filter are considered true.

2.6.11 Statistical analysis

All experiments include at least three independent biological replicates and results are presented as the mean value (± SD). Statistical analyses were conducted using a Welch two sample *t*-test performed in R[158]. The significance level was set at 0.05 for all statistical analyses.

2.7 Acknowledgments

We are grateful to Dr. Alicia Withrow of the MSU Center for Advanced Microscopy for her extensive assistance with the Transmission Electron Microscope and Dr. Xudong Fan for assistance with EDX analyses. The Ccm antibodies used in this study were graciously provided by Dr. Cheryl Kerfeld. We also thank Dr. Kerfeld and members of her research group, as well as Dr. Danny Ducat and members of his research group, for detailed discussions about experimental approaches and results from our analyses.

This work was supported by the U.S. Department of Energy (Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, grant no. DE-FG02-91ER20021 to B.L.M.) and support to Shailendra Singh from the National Science Foundation (grant no. MCB-1243983 to B.L.M.). We also thank Melissa Whitaker (supported by National Science Foundation grant no. MCB-1243983 to B.L.M.) for strain maintenance and culture production.

CHAPTER 3

Cyanobactarial Carboxysomes Contain a Unique Rubisco-Activase-Like Protein

This chapter contains information published in:

Lechno-Yossef, S., Rohnke, B. A., Belza, A. C., Melnicki, M. R., Montgomery, B. L., and Kerfeld, C. A. (2019) Cyanobacterial carboxysomes contain a unique rubisco-activase-like protein. *New Phytol. early view*. doi: 10.1111/nph.16195

reproduced with permission and modified to incorporate the supplemental information into the body of the text, reorder subsection **3.4 (**Materials and methods) to be after subsection **3.3** (Discussion), renumber figures, tables, and references to be consistent with the dissertation, and use abbreviations defined in the KEY TO ABBREVIATIONS.

Brandon Rohnke was the primary contributor to sections **3.2.2**, **3.2.5**, **3.2.6**, **3.4.2**, **3.4.11**, & **3.4.12**, Tables 3.2 & 3.3, and Figures 3.3, 3.5, 3.10, 3.11, & 3.12. He contributed in part to sections **3.1**, **3.2.1**, **3.3**, & **3.4.1**, Figure 3.2, and editing of the overall manuscript.

3.1 Introduction

Rubisco is the most abundant enzyme in nature and the major enzyme responsible for primary productivity on earth[52]. The most prevalent type, Form I, is composed of large and small subunits arranged in an L_8S_8 configuration, and is found in all photosynthetic eukaryotes and β-cyanobacteria (Form IB), as well as in all α-cyanobacteria and certain anoxygenic phototrophs (Form IA). The enzyme has a low catalytic turnover rate, suffers from energetically wasteful side-reactions with O_2, and is prone to forming inactive complexes upon binding sugar phosphates, including its own substrate, RuBP. In order to relieve rubisco from these inactive states, plants utilize the molecular chaperone rubisco activase (Rca)[152]. This enzyme has been the focus of much study, including as part of biotechnological efforts to develop strategies to improve rubisco performance[22,131].

Rca belongs to the AAA+ family of proteins (ATPases associated with diverse cellular activities), of which many use the energy from ATP hydrolysis to remodel the conformation of another protein[136]. Accordingly, Rca removes inhibitory sugar phosphate effectors from rubisco[172] by coupling ATP hydrolysis with destabilization of the rubisco active site, in a yet-unknown mechanism[27]. Crystal structures of Rca from tobacco (*Nicotiana tabacum*) and *Arabidopsis thaliana* exhibit critical features of the AAA+ fold, including its tendency to oligomerize in hexameric and dynamic helical states[73,196]. [ATP]-dependent variability of oligomeric state[99,149,215] and presence of a redox-sensitive C-terminal extension contribute to the regulation of plant Rca activity by conditions such as light intensity and CO_2 concentration[152,223]. Rca deletion in plants results in an HCR phenotype, which suggests a critical role for the activase in maintaining plants' capacity for carbon fixation[191].

By contrast, inactivation of rubisco by sugar phosphates is less clear in cyanobacteria, with most information obtained from organisms that lack the activase-like-protein of cyanobacteria (ALC). Unlike the plant enzyme, cyanobacterial rubisco exhibits a lower affinity for CO_2 and RuBP, lower CO_2/O_2 specificity, and higher maximal carboxylation rate[5,147,163,207]. However, differences in inhibition of the uncarbamylated enzyme by RuBP have been noted between species. No inhibition by RuBP was observed in unicellular cyanobacteria[5,147,163], whereas inhibition with RuBP was reported in the filamentous heterocyst-forming *Anabaena variabilis*, but with a K_i value curiously similar to the K_m measured for RuBP carboxylation[9] and an absence of significant inhibition when testing recombinant rubisco *in vitro*[102]. From these disparate studies, it is difficult to characterize whether cyanobacterial rubisco is indeed prone to inhibition by RuBP or other sugar phosphates, and moreover whether an activase is required. Indeed, until genome sequences from diverse cyanobacteria became available, it was believed that only very few cyanobacteria contained an ALC[107].

The lower specificity of the cyanobacterial rubisco to CO_2 is likely the result of its compartmentalization in carboxysomes as part of a CCM[14,86,91]. These bacterial microcompartments encapsulate rubisco and CA within a polyhedral protein shell, which is believed to maintain an internally-elevated CO_2 concentration[34]. The eco-physiologically diverse β-cyanobacteria employ β-carboxysomes, encapsulating Form IB rubisco and an absolutely-conserved γ-class CA named CcmM[160,192]. CcmM was shown to aggregate rubisco for β-carboxysome assembly; it contains 3-5 copies of SSLDs, accessory domains with sequence homology to RbcS[38,112,142]. CcmM has two alternative start sites resulting in two versions of the

protein: the full-length M58 contains both the CA domain and the SSLDs, while a shorter version – M35 – contains only the SSLDs[114]. The structure of a single SSLD derived from a CcmM contains an additional alpha helix (helix H1A) not found in RbcS[171,213]. Additionally, the corresponding M35 was shown to be involved in rubisco nucleation and network formation in a redox-sensitive manner, interacting with the rubisco L_8S_8 complex equatorially rather than by replacement of an RbcS subunit with one of its SSLDs, as previously predicted[213].

Given their absence from the most commonly used cyanobacterial laboratory strains, ALCs have received scant attention. Moreover, almost twenty years ago, an ALC deletion mutant in *Anabaena variabilis* showed only a mild reduction of growth rate in air[103]. However, with the increased availability of cyanobacterial genomic sequence data, it became apparent that many β-cyanobacteria encode Rca-like genes, consisting of the AAA+ ATPase domain fused to a single C-terminal SSLD[222]. These ALCs have been proposed to be ancestral to the plant Rca[135].

Here, we phylogenomically survey ALCs and reveal that they are widespread among β-cyanobacteria but are absent from one major taxonomic subclade of cyanobacteria which is predominated by unicellular species and includes several common lab strains. We found a correlation between ALC and RbcL phylogenetic subtypes encoded in the same genome, suggesting co-evolution between the ALC and RbcL. Structural modeling also suggests conservation of ATPase activity, hexameric assembly, and potential interaction between ALC and rubisco. We biochemically and physiologically characterized the ALC from *F. diplosiphon* UTEX 481 (also known as *Tolypothrix* sp. PCC 7601 and *Calothrix* sp. PCC 7601), demonstrating ALC localization to carboxysomes and its close physical proximity to rubisco

encapsulated there. We establish that *F. diplosiphon* ALC (*Fd*ALC) is active as an ATPase, but does not relieve inhibition by RuBP, nor does *alc* deletion result in obvious growth defect. Instead we find that ALC induces rubisco aggregation in an ATP-dependent manner, and that its deletion affects regulation of carboxysome biogenesis and cell morphology in response to CO_2 levels. ALC is not a canonical activase, instead exerting its effect on CO_2 fixation at the level of a metabolic module, the carboxysome.

3.2 Results

3.2.1 Bioinformatic analysis of the ALC gene family

The ALC protein sequences were identified in cyanobacterial genomes in the Integrated Microbial Genomes database (IMG; https://img.jgi.doe.gov/) (Table S1) based on presence of the AAA+ ATPase domain (pfam00004) and each of the three highly-conserved "non-canonical pore loop" regions defined previously for Rca proteins[196]. The AAA+ domains of the resulting 133 sequences (Table S1) were aligned together with AAA+ domains from reference *rca* sequences from plants and algae and subsequently used to build a phylogenetic tree (Figure 3.1), within which six clusters could be manually identified, representing groups with shared sequence similarity. Eight ALC sequences could not be assigned to any of the six clusters (Marked ALC-X in Table S1). All but 10 sequences also contained a C-terminal extension encoding an SSLD (Table S1). The comparison between the predicted cyanobacterial ALC structure and that of plant Rca is discussed in subsection 3.2.2. No genome encodes more than one ALC, thus suggesting a lack of duplication and therefore orthology among the ALCs.

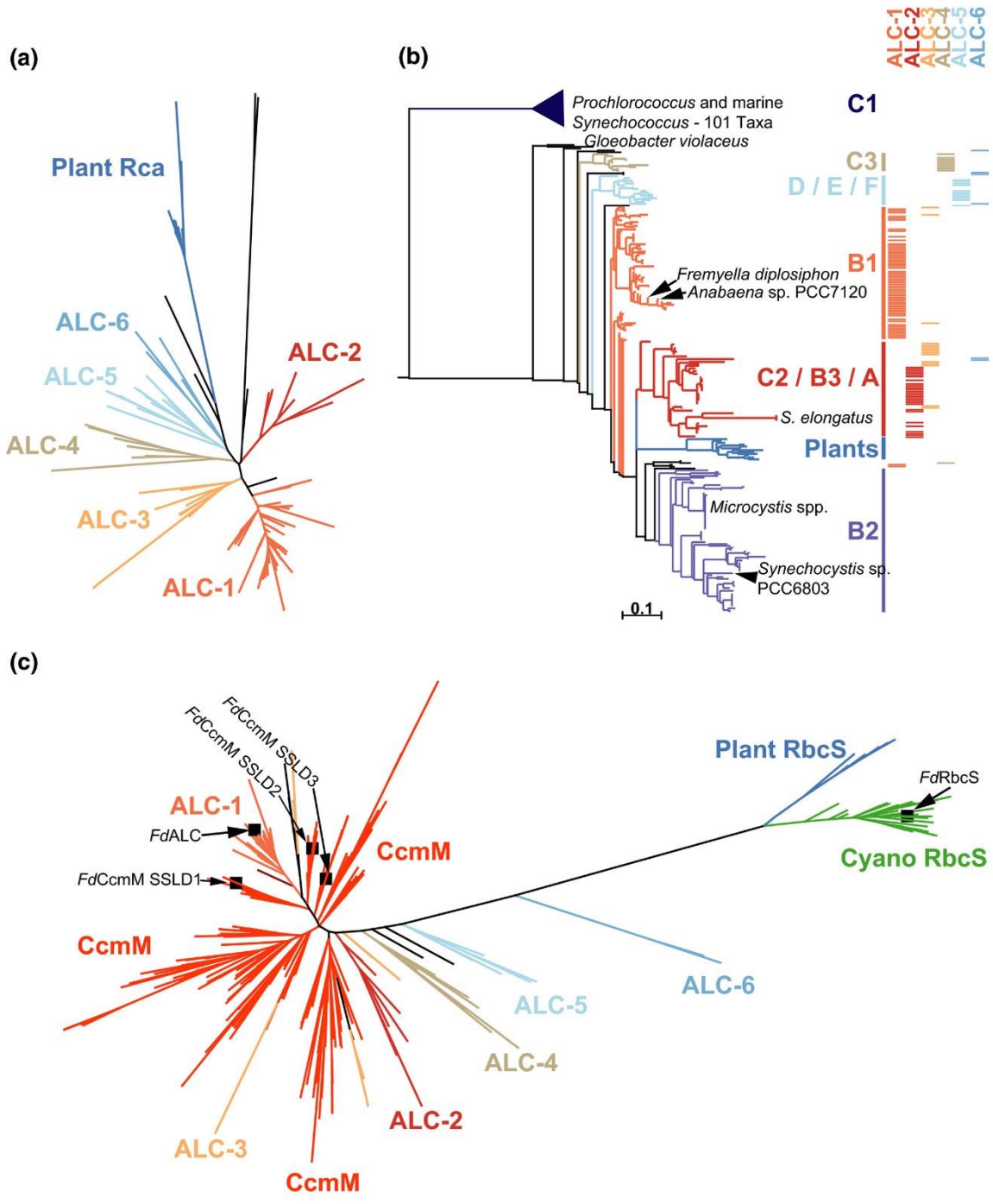


Figure 3.1: Phylogenetic analysis of cyanobacterial ALC and RbcL sequences. (A) Unrooted phylogeny of AAA+ domains of ALCs and Rca. The AAA+ domain of 133 ALC sequences were identified in 374 cyanobacterial genomes, aligned with 10 plant and algal Rca reference sequences (Table S1), and used for phylogenetic tree construction. **(B)** Phylogeny of Form IB RbcL and correspondence with ALC subtypes. RbcL sequences from 335 cyanobacterial genomes were aligned with 13 plant and algal representatives (Table S1), used for phylogenetic tree construction, and were rooted by the Form IA rubisco sequences from α-cyanobacteria (subclade C1, shown collapsed). Vertical bars depict the phylogenetic subclades of the host genomes (based on Shih et al.[185]). Columns at right show correspondence with ALC subtypes when present in the same genome. **(C)** Unrooted phylogenetic tree comparing SSLDs from CcmM (179 domains in 51 CcmM sequences) and ALC (56 representative sequences) with RbcS (56 representative sequences).

In order to examine the potential interaction and coevolution between rubisco and ALC, cyanobacterial RbcL sequences were aligned with the RbcL sequences from the 13 reference plant and algal species (Table S1). Phylogenetic analysis shows that RbcL sequences generally cluster together by the taxonomic groups ("phylogenetic subclades") previously determined phylogenomically by Shih et al.[16,185], although the branching order between the subclades is different, and some subclades could not be resolved based on RbcL alone (Figure 3.1B). When ALC clade IDs are mapped to the branches of RbcL sequences corresponding to the same genome, it can be seen that all species from subclade B2 lack an ALC. The organisms in this subclade are dominated by the unicellular *Chroococcales*, and include several model organisms such as *Syn*6803, *Synechococcus* PCC 7002, and *Cyanothece* ATCC 51142, as well as other well-studied strains belonging to *Microcystis* and *Crocosphaera*. The ALCs are otherwise distributed widely across all other β-cyanobacteria, with a general correlation between taxonomic subclade, RbcL sequence, and the type of ALC present (Table 3.1). Most notably, the RbcL sequences of heterocyst-forming cyanobacteria from morphological subsections IV and V[165] – or subclade B1 – cluster together in one region of the tree, and mostly contain ALC-1. Similarly, the filamentous non-heterocyst-forming species (morphological subsection III) from taxonomic subclade A correspond with ALC-2. The same RbcL clade, however, was intermixed with organisms from taxonomic subclade C2, which lack an ALC, and subclade B3, which generally contain ALC-3. RbcL sequences from organisms in subclade C3 (subsections I and III) clustered in a monophyletic clade and are associated with ALC-4. RbcL sequences from organisms belonging to subclades D, E and F (subsections I and III) also clustered together and were found to contain either ALC-5 or ALC-6 (Figure 3.1B), which were the ALC types branching closest to the plant Rca cluster (Figure 3.1A); in contrast, the plant RbcL sequences are located quite

Table 3.1: Correlation between ALC subtype and taxonomic subclade. Reference subclade and subsection names are as published[165,185].

	# members	Shih subclade[a]	Rippka subsection[a]
ALC-1	60	B1	IV, V, (II)
ALC-2	23	A	III
ALC-3	16	(B1), B3, C2	I, II, III
ALC-4	10	C3, (G)	I, III
ALC-5	10	(A), D, E	I, III
ALC-6	6	E, F	I, III
ALC-X	8	Various	Various

[a] A subclade or subsection in parentheses denotes a single sequence belonging to that group.

distant from the D/E/F RbcL subcluster. Overall, ALCs show strong sequence similarity to Rca from green-lineage photosynthetic organisms (62% pairwise identity between consensus sequences for the ALC vs Rca AAA+ domains (Table 3.2).

Table 3.2: Pairwise sequence alignment scores of ALC and RbcL sequences. Scores were generated using the LAlign webserver. Consensus sequences were generated in CLC Sequence Viewer from the respective MSA and had no gaps.

Protein	Sequence 1	Sequence 2	Percent Sequence Identity (%)	Percent Sequence Similarity (%)
ALC	Consensus of Cyanobacteria	Consensus of Higher-Order Green-Lineage	62.2	85.1
	F. diplosiphon	*A. thaliana*	62.5	85.1
	F. diplosiphon	*N. tabacum*	61.0	86.4
	N. tabacum	*A. thaliana*	88.1	95.8
Form IB RbcL	Consensus of Cyanobacteria with ALC	Consensus of Higher-Order Green-Lineage	86.7	96.6
	Consensus of Cyanobacteria with ALC	Consensus of Cyanobacteria without ALC	90.8	97.6
	Consensus of Cyanobacteria without ALC	Consensus of Higher-Order Green-Lineage	85.4	95.9
	F. diplosiphon	*A. thaliana*	83.2	95.3
	F. diplosiphon	*N. tabacum*	83.3	95.3
	N. tabacum	*A. thaliana*	94.3	98.1
	F. diplosiphon	*S. elongatus*	85.2	96.6

A phylogenetic tree built with RbcS and SSLDs (from either ALC or CcmM) retains the same ALC subclustering and branch ordering as the AAA+-based tree (Figure 3.1C). However, the CcmM-derived SSLDs are interspersed among those from ALC-3, representing a high degree of sequence similarity among all SSLDs, supported by conservation of characteristic primary and secondary structural features (Figure 3.2) previously described for CcmM[116]. Because SSLDs from CcmM have been shown to interact with rubisco[38,44,112,213], the similarity shown here among all SSLDs suggests that those derived from the ALC are likely to interact with rubisco in the same manner.

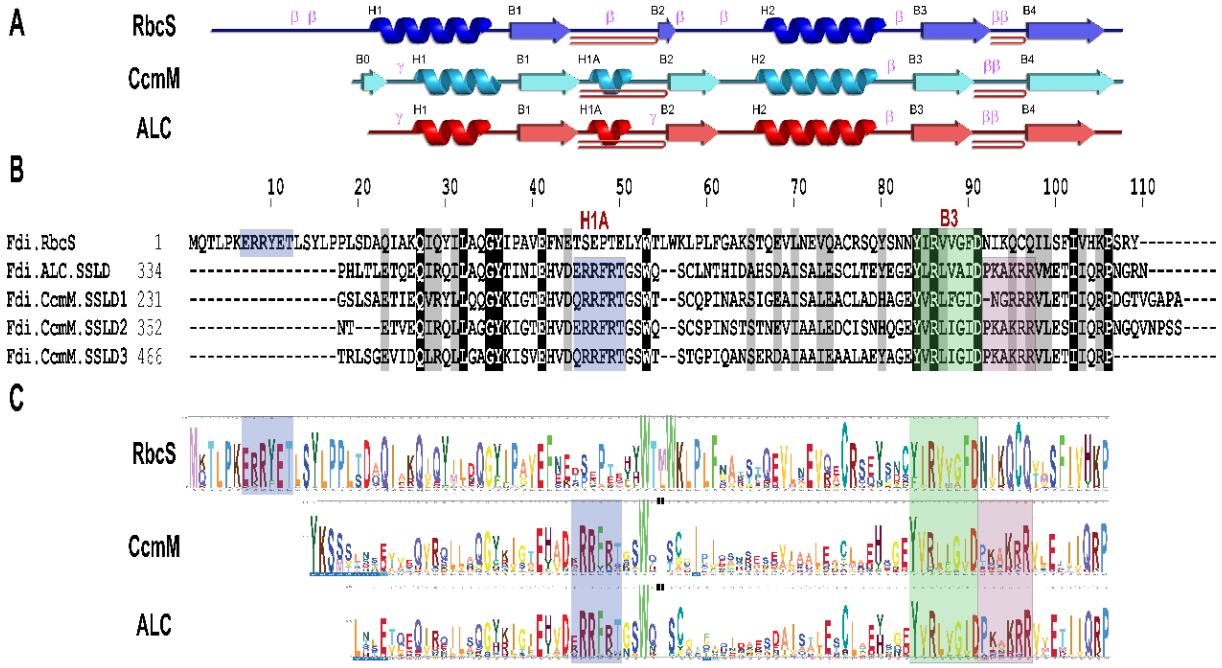


Figure 3.2: Primary and secondary structure comparison of RbcS and the SSLDs of *Fd*ALC and CcmM. (A) Schematic of the secondary structure for RbcS from *Synechococcus elongatus* PCC 6301 (hereafter *Syn*6301) (blue, PDB: 1RBL), CcmM-SSLD from *T. elongatus* (turquoise, PDB: 6MR1), or a homology model of *Fd*ALC-SSLD (red) generated in PDBsum. Helices are labeled H1-H2, β-strands are labeled B0-4. β, β-turns; γ, γ-turns; red horseshoe, β-hairpins. **(B)** MSA of *F. diplosiphon* RbcS, *Fd*ALC-SSLD (residues 334 – 424), and three *F. diplosiphon* CcmM SSLDs (residues 231 – 324, 353 – 447 and 353 – 447). Black highlights, identical residues in all sequences; grey highlights, similar residues; pale blue highlight, conserved RbcS N-terminal domain with the motif ERRYET, has high sequence similarity to a motif found in the helix H1A of the SSLD (E/QRRFRT); green highlights, conserved β-strand 3, pink highlights, a loop region which is conserved only in SSLDs and not in RbcS. **(C)** Sequence logos of RbcS sequences from 223 β cyanobacteria, 111 CcmM-SSLD domains from 32 β-cyanobacteria and 108 ALC-SSLD sequences.

3.2.2 Structural modeling of the ALC

In order to compare the ALC to canonical plant Rca, we modeled the ALC from *F. diplosiphon* (hereafter: *Fd*ALC) as a monomer, using 3D crystal structures for Rca from tobacco (PDB: 3T15) and an SSLD from *Thermosynechococcus elongatus* CcmM (PDB: 6MR1) as templates. *Fd*ALC residues 1 – 293 were selected as the target sequence for the Rca template, and residues 334 – 424 were used for the SSLD template. Pairwise alignments of *Fd*ALC with Rca and CcmM-SSLD had 86.5% and 80.0% sequence similarity respectively, suggesting a well-conserved AAA+ domain with a flexibly-positioned SSLD extending C-terminally from helix H10 (Figure 3.3A & 3.3B). Because Rca and other AAA+ ATPase proteins function as hexamers, we also predicted the ALC quaternary structure. As shown previously with monomeric Rca[196], our monomeric ALC model includes some domain shifts relative to a hexameric model for Rca (Figure 3.3C). Thus, the quaternary structure of *Fd*ALC was also modeled using the hexameric model from tobacco as a template (Figure 3.3D). A hexameric assembly of the ALC is also suggested by size exclusion chromatography and dynamic light scattering analysis of the purified protein (Figure 3.4).

In order to study the structural conservation of the ALC, we mapped the conservation scores for each position in an ALC sequence alignment on to the full-length *Fd*ALC homology models. Patches of strong sequence conservation were found to cluster into sectors within the monomeric 3D model (Figure 3.5A), particularly in a region corresponding to what is conventionally deemed the top side of Rca and related AAA+ hexamers, as previously observed with plant Rca[196]. When the *Fd*ALC is viewed as a hexamer, the top side appears to be almost completely conserved (Figure 3.5B), whereas the bottom and perimeter regions exhibit more sequence

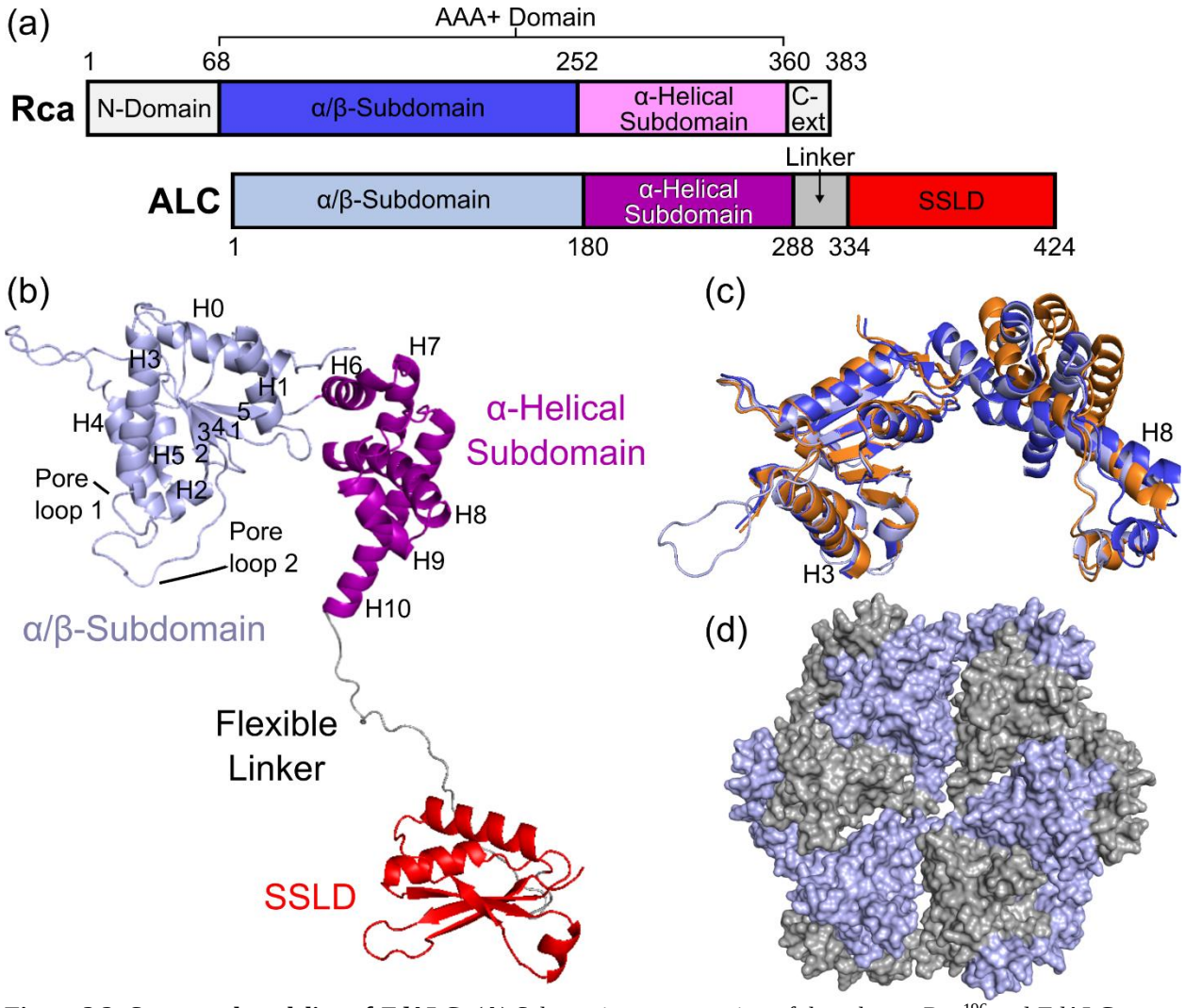


Figure 3.3: Structural modeling of *Fd*ALC. (A) Schematic representation of the tobacco Rca[196] and *Fd*ALC domains. C-ext, C-extension; SSLD, Small Subunit-Like Domain. **(B)** Homology model of full-length *Fd*ALC using tobacco Rca and *T. elongatus* CcmM-SSLD as templates for the AAA+ domain and SSLD, respectively. For the AAA+ domain, helices are labeled H0-H10, β-strands are labeled 1-5. **(C)** Alignment of the *Fd*ALC AAA+ domain homology model (light blue), tobacco Rca (dark blue), and tobacco Rca modeled to fit a hexameric AAA+ domain (orange). **(D)** Homology model of *Fd*ALC hexamer using the modeled tobacco Rca hexamer as a template, viewed from the top side. Monomeric subunits alternate between grey and light blue coloration.

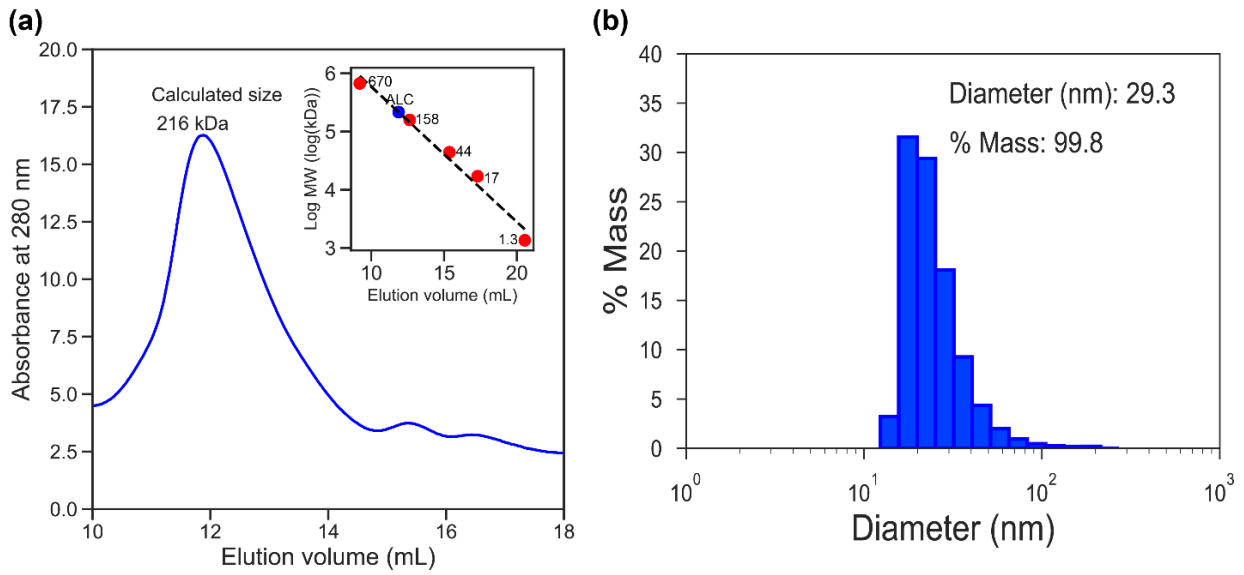


Figure 3.4: The oligomeric state of purified recombinant *Fd*ALC. (A) Size exclusion chromatography of *Fd*ALC resolved on Superdex S200 GL10/30 column. Resolution of standards of known molecular weight on the same column is shown in the inset. The *Fd*ALC is eluted in a wide non-symmetrical peak suggesting a mixture of several oligomeric states, and absence of a monomeric form. Similar profile was obtained when tobacco Rca was resolved by SEC[196]. **(B)** Dynamic Light Scattering (DLS) of purified *Fd*ALC. (DynoPro Nanostar by Wyatt Technology, Santa Barbara, CA was used for the analysis). The majority of analysed structures had an average diameter of 29.3 nm. The hexameric model (Figure 3.3D) has a diameter of 13 nm, but does not include the dynamic SSLD, hence a size of 29.3 nm, could potentially correspond to a hexamer, and is certainly larger than a monomer.

variability (Figure 3.5C & 3.5D). Additionally, as the SSLD domain extends from the C-terminus of each ALC subunit, it was modeled pointing toward the bottom side of the hexamer; however, the linker is likely flexible and thus the SSLD may possibly adopt one or more alternative configurations. Nevertheless, in agreement with the nearby variability of the bottom side of the hexamer, both the linker and the SSLD show low sequence conservation, with the exception of one well-conserved region in the SSLD. In the AAA+ domain, the α/β subdomain contains much greater sequence conservation than the α-helical subdomain (Figure 3.5A), which is consistent with plant Rca[196], as well as the functional role that the α/β subdomain plays in ATP binding. In particular, residues implicated in ATPase and/or activase function[196] show very high conservation scores (Figure 3.5E).

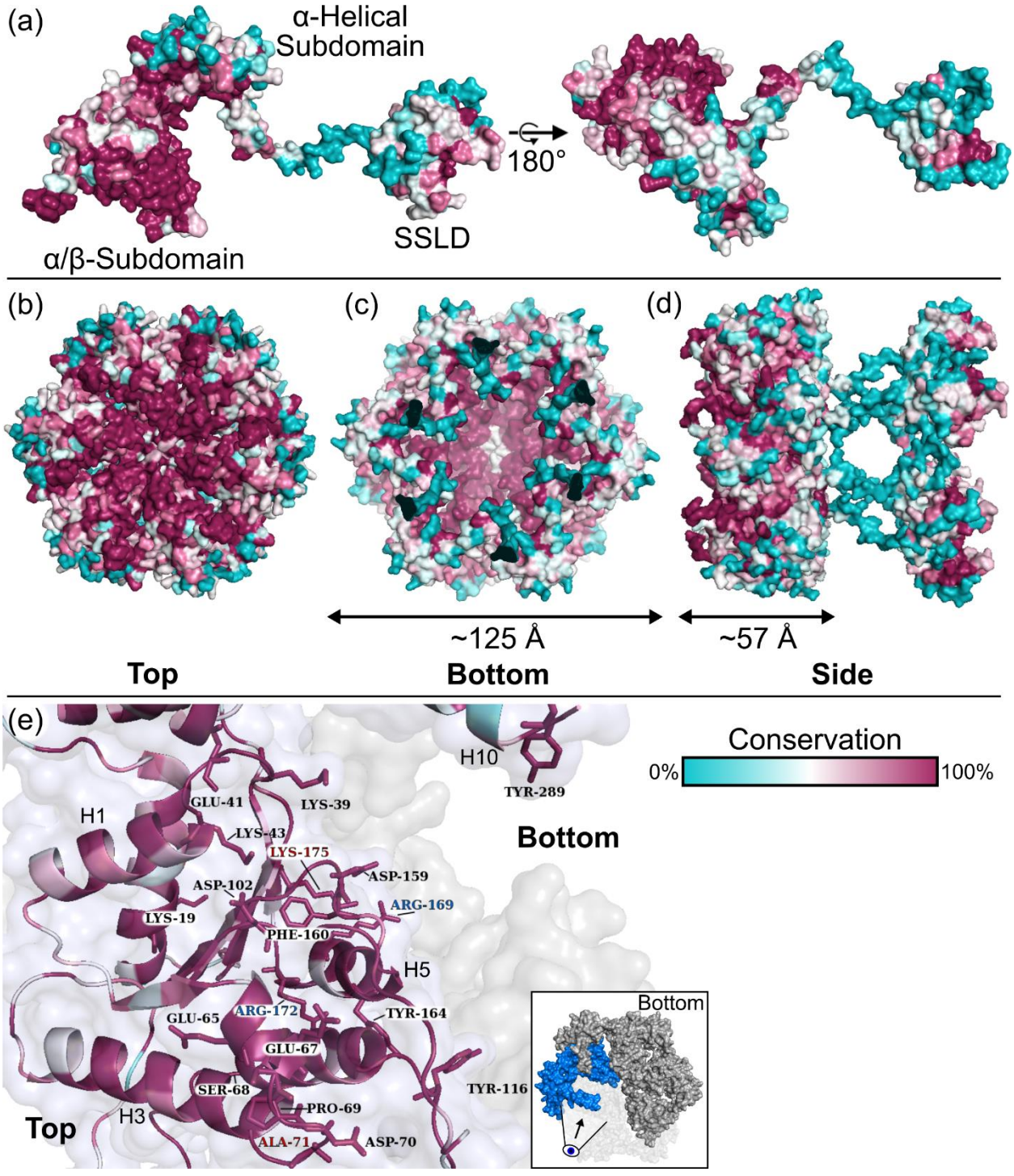


Figure 3.5: *Fd*ALC evolutionary conservation structural models. Conservation scores based on an MSA of 133 ALC sequences mapped onto the *Fd*ALC homology model. **(A)** Surface representation of the monomeric form. **(B)** Top, **(C)** bottom (SSLDs were removed from the plane of view), and **(D)** side views of the *Fd*ALC modeled as a hexamer (Figure 3.3D). **(E)** Residues of known importance for activase (red labels) and ATPase (black labels, with arginine fingers labeled in blue) function as summarized by Stotz et al.[196]. View is indicated by the arrow pointing towards the blue colored monomer in the inset.

3.2.3 Cellular localization of the ALC from *F. diplosiphon*

If the ALC functions similarly to Rca, it would require proximity to its substrate, rubisco, which is encapsulated inside carboxysomes. To test whether the ALC is localized to carboxysomes, the fluorescent protein mTurquoise2 (mT2) was translationally fused to the C-terminus of full-length *Fd*ALC. This construct was expressed in *F. diplosiphon* cells containing a second construct in which the Yellow Fluorescent Protein (YFP) was fused to the N-terminus of RbcL. When co-expressed, the fluorescence patterns from both *Fd*ALC-mT2 and YFP-RbcL show puncta that are characteristic of carboxysome localization[38,128,174,199], with both blue and yellow fluorescence channels appearing in overlapping patterns throughout the cells, suggesting that ALC and RbcL proteins are co-localized and thus implying that the ALC is targeted to carboxysomes (Figure 3.6A).

Förster resonance energy transfer (FRET)[87] was used to further elucidate potential interactions between *Fd*ALC-mT2 and YFP-RbcL. After an initial image was collected (Figure 3.6A), the YFP signal in a selected region was intentionally photobleached (using the microscope laser), and a second image of the same field was taken (Figure 3.6B). An increased mT2 fluorescence after photobleaching indicates FRET between the mT2 and YFP (Figure 3.6C). This positive detection of FRET between *Fd*ALC-mT2 and YFP-RbcL suggests that the two proteins must be within 10 nm of each other inside the carboxysome, which has an average diameter of about 360 nm in *F. diplosiphon*, as measured by TEM (Section **3.2.6**; Table 3.3).

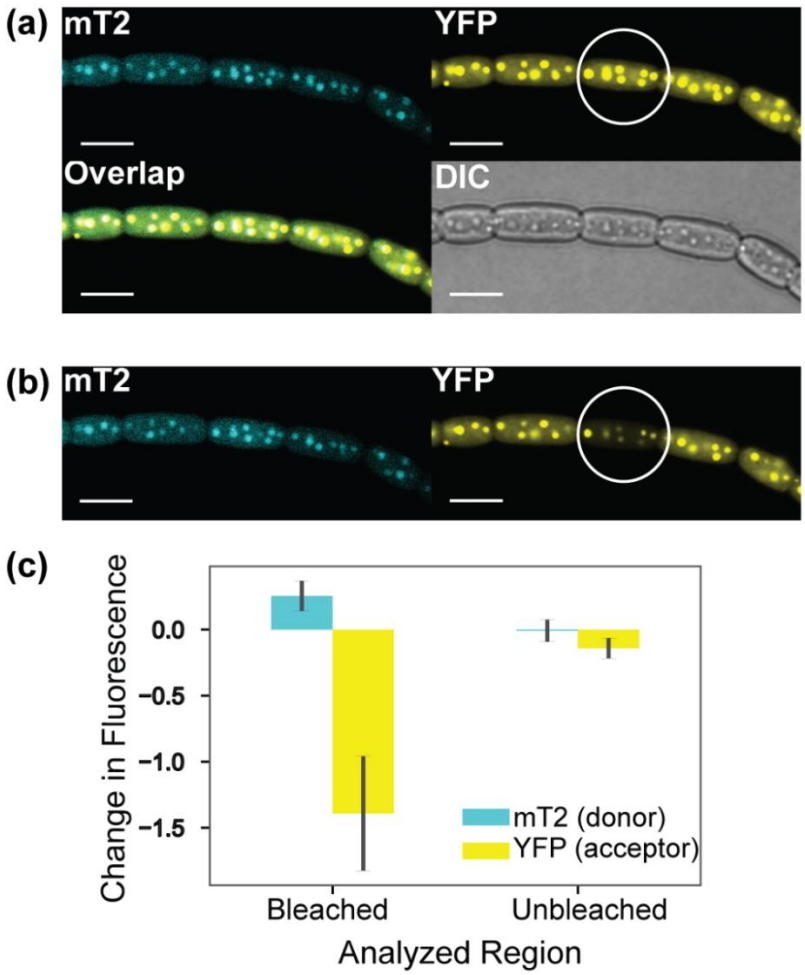


Figure 3.6: Co-localization of ALC and RbcL in carboxysomes of *F. diplosiphon*. Confocal microscopy was performed on a strain co-expressing plasmids carrying the *Fd*ALC-mT2 fusion and another carrying the YFP-RbcL fusion. A representative field of view is presented for blue channel fluorescence (mT2), yellow channel fluorescence (YFP), overlap of the blue and yellow channels, as well as the corresponding differential interference contrast. Scale bars represent 5 μm. **(A)** Image taken before photobleaching. **(B)** Blue and yellow channel fluorescence of the field shown in (A), after photobleaching of the yellow signal in a targeted region indicated by a white circle. **(C)** Calculated change in blue and yellow fluorescence before and after photobleaching. Averages from 10 imaged fields are shown.

3.2.4 The *F. diplosiphon* ALC has ATPase activity and interacts with rubisco, but does not function as rubisco activase

Given that *Fd*ALC contains the residues essential for ATPase activity (Figure 3.5E), we investigated its activity upon purification as a C-terminal His tag fusion and expression in *E. coli* (Figure 3.7A). Using an enzyme-linked NADH oxidation assay[132,202], ATPase activity in *Fd*ALC was confirmed, with a maximum ATP hydrolysis rate (V_{max}) of 349 ± 60 μmol mg^{-1} min^{-1}

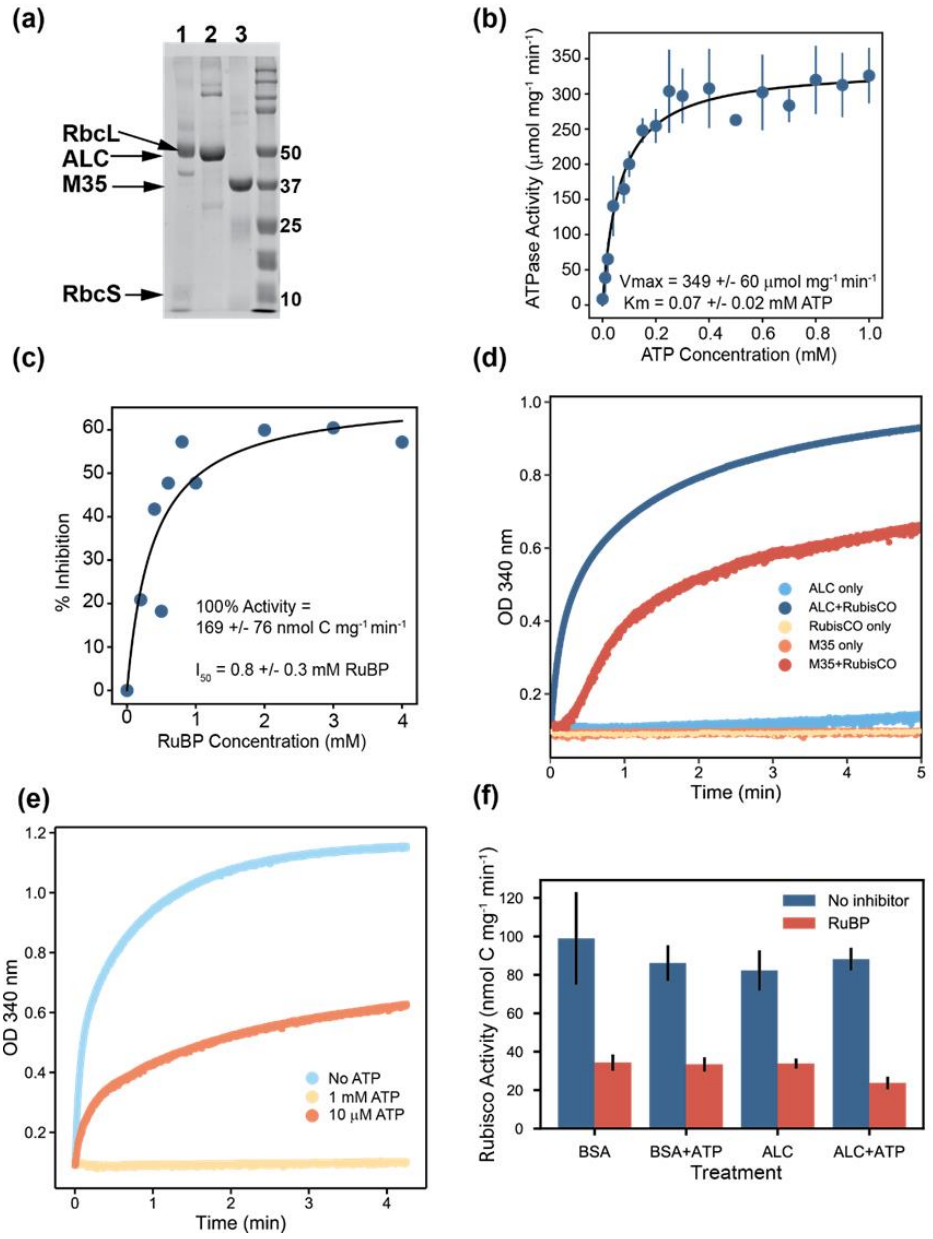


Figure 3.7: Biochemical activity of recombinant rubisco and *Fd*ALC. (A) Coomassie-stained SDS-PAGE of proteins purified from *E. coli* by affinity chromatography used for biochemical assays. Lane 1 – rubisco, Lane 2 – *Fd*ALC, Lane 3 – M35. **(B)** Michaelis-Menten kinetics ATP hydrolysis by *Fd*ALC expressed and purified from *E. coli*, using 124 µg protein mL^{-1} with varying concentrations of ATP (n = 3). **(C)** Effect of RuBP concentration on rubisco activity. *F. diplosiphon* rubisco was pre-incubated with different concentrations of RuBP before activation with $MgCl_2$ and bicarbonate followed by incorporation of radioactive bicarbonate (n = 3) **(D)** Kinetics of rubisco aggregation mediated by the *Fd*ALC and the M35. Change of turbidity at 340 nm was followed after adding rubisco with or without *Fd*ALC or M35 to the assay. Representative experiment is shown (n = 3). **(E)** The effect of ATP on ALC-mediated rubisco aggregation. Assay performed as described in (D), with the addition of creatine phosphate, creatine phosphokinase and ATP at stated concentrations to the reaction buffer. Representative experiment is shown (n = 3). **(F)** Effect of the *Fd*ALC on rubisco activity. Rubisco activity following incubation with (green bars) or without (blue bars) RuBP was assayed as in (B), activation step was conducted as described[17]. The presence of *Fd*ALC or bovine serum albumin (as a negative control), and/or 5 mM ATP is stated. Data from a representative biological replicate with 3 technical replicates each are shown (experiment was repeated with two more biological replicates).

(Figure 3.7B), which is higher than the V_{max} reported for spinach Rca[167], 1.5 μmol mg^{-1} min^{-1}. Unlike the spinach Rca, the Michaelis-Menten plot for *Fd*ALC is hyperbolic rather than sigmoidal (Figure 3.7B, compared to Figure 3 in Robinson & Portis[167]), suggesting that under the experimental conditions tested the ATP binding in *Fd*ALC may not be cooperative. The calculated K_m for *Fd*ALC is 0.07 ± 0.02 mM ATP, comparable to the K_m for spinach Rca[167], implying that the ALC enzyme might operate at a similar range of physiological ATP concentrations.

In order to investigate whether *Fd*ALC can affect the activity of its cognate rubisco, we first purified recombinant *F. diplosiphon* rubisco that was expressed in *E. coli* using a three-plasmid system modified from Aigner et al.[2]. Rubisco was purified by affinity chromatography, yielding a complex containing large and small subunits (Figure 3.7A), and activity was confirmed showing incorporation of radioactive bicarbonate.

In contrast to previous studies showing a lack of RuBP inhibition for cyanobacterial rubisco from various species[5,102,147,163], here we have detected inhibition of recombinant *F. diplosiphon* rubisco carbamylation by RuBP, reaching a calculated maximum of 85% inhibition, with half-maximum of the inhibition occurring at 0.8 ± 0.3 mM RuBP (Figure 3.7C). While variability among the biological replicates (that may be due to variability among the protein preps derived from the *E. coli* expression system) precludes estimating precise values for inhibition parameters, it is clear that RuBP inhibits *F. diplosiphon* rubisco. Moreover, rubisco purified directly from *F. diplosiphon*, likewise is clearly inhibited by RuBP (Figure 3.8).

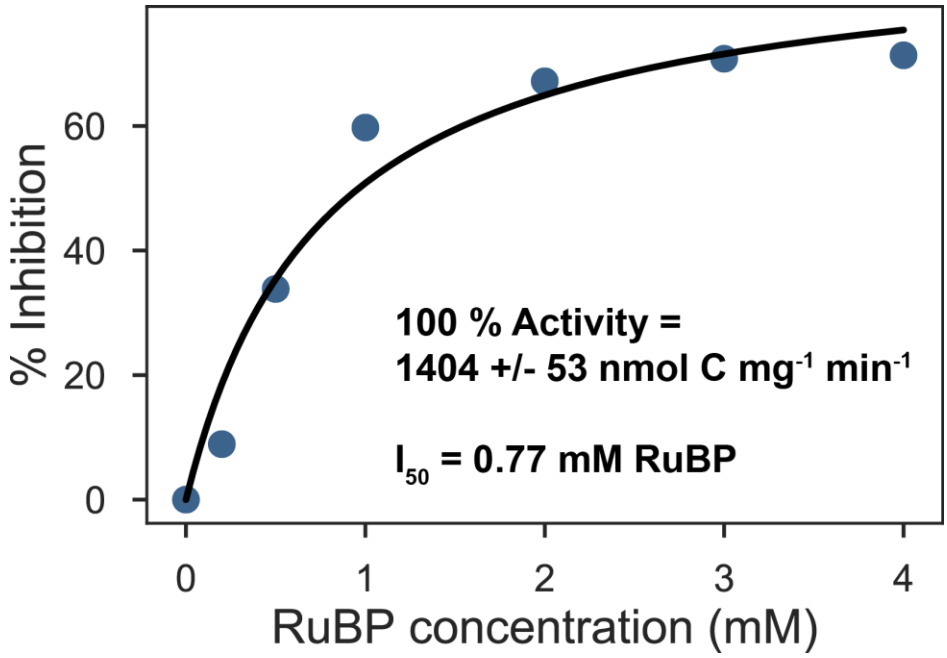


Figure 3.8: Effect of RuBP concentration on activity of rubisco purified from WT *F. diplosiphon*. Rubisco was purified from *F. diplosiphon* by differential ammonium sulfate precipitation and ion exchange chromatography. Assay was conducted as described for Figure 3.7C. Average of 3 technical replicates is shown.

Because the M35 and the ALC both contain SSLD domains, the effect of the ALC on rubisco was compared to that of recombinant M35 purified from *E. coli*. Similar to the recently-shown effect of M35[213], addition of *Fd*ALC to rubisco causes an increase of turbidity, suggesting rubisco aggregation (Figure 3.7D). ALC-mediated rubisco aggregation took place only if the intact ALC was used, and not when either of the two separate domains (AAA+ domain or the SSLD), prepared as two synthetic proteins were mixed with rubisco, emphasizing the importance of both domains in interaction with rubisco (Figure 3.9). This ALC-mediated effect was affected by the presence of ATP (Figure 3.7E). Only a slight increase in turbidity was observed when the ALC was incubated in the absence of rubisco (Figure 3.7D). Rubisco activity was unaffected by the presence of rubisco aggregates (Figure 3.7F, blue bars), and unlike the effect of Rca on plant rubisco, the presence of recombinant *Fd*ALC with ATP did not restore activity to rubisco preparations that were pre-inhibited by 4 mM RuBP (Figure 3.7F, green bars), thus indicating that *Fd*ALC does not display canonical activase activity under the conditions tested.

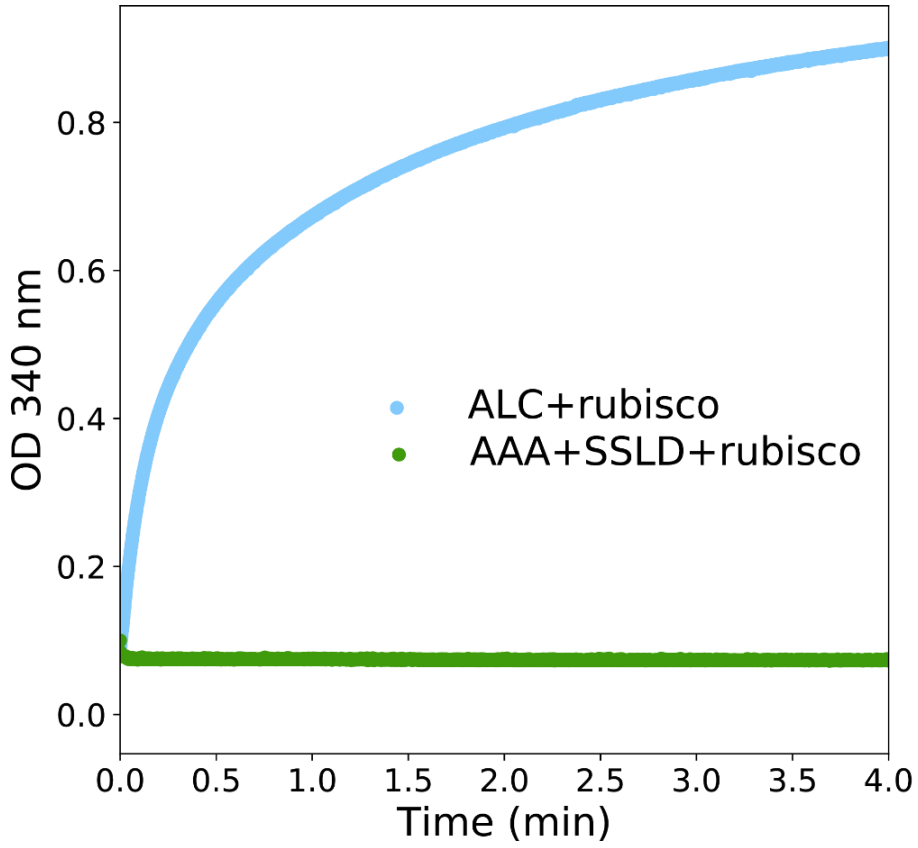


Figure 3.9: The effect of full length ALC compared to its two separate domains on rubisco aggregation. Rubisco aggregation was tested by mixing 0.25 µM recombinant rubisco with either 2 µM *Fd*ALC (blue line) or 6 µM SSLD and 2 µM AAA+ domain (green line) purified separately from pDS2 and pSL283, respectively, as described[101]. Assay conditions as described for Figure 3.7. Representative result shown (n = 3).

3.2.5 *F. diplosiphon* ALC is upregulated under CO_2 enrichment

To confirm whether *FdALC* is expressed *in vivo*, we examined transcript levels using quantitative qPCR. Transcript levels of *alc* and *ccmM* in response to increasing carbon availability were measured. The expression levels of *alc* were an average of ~4-fold higher in WT strains grown under air enriched with 3% CO_2, compared to C_i-limited cells that had been transferred from CO_2 enrichment to air in order to induce C_i-stress from CO_2 downshift. Additionally, expression levels also increased upon CO_2-upshift for *ccmM*, which is known to impact carboxysome size and abundance[112,170], but this response was dampened in a Δ*alc* mutant strain that was generated to test for a functional role of *Fd*ALC (Figure 3.10).

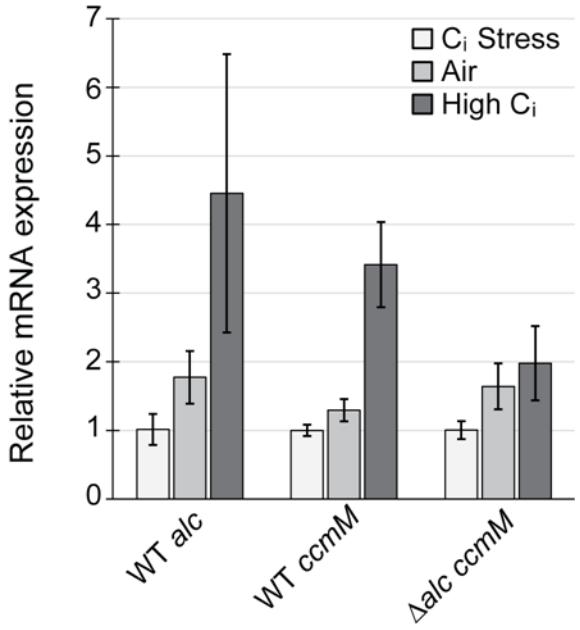


Figure 3.10: Relative expression of *alc* and *ccmM* genes under varying C_i levels. qPCR data of WT and Δ*alc* mutant strain of *F. diplosiphon* grown with (high C_i) or without (Air) enrichment of 3% CO_2. Some samples grown with CO_2 enrichment were transferred to air 18 hr prior to harvesting in order to induce low C_i stress (C_i Stress). Data are expressed based on $\Delta\Delta C_T$ analysis using C_i Stress as the reference condition. Bars, 95% confidence interval.

3.2.6 Ultrastructural characterization of a *F. diplosiphon* ALC deletion mutant shows misregulated response to CO_2 availability

Similar to prior results for *Anabaena variabilis*[103], a *F. diplosiphon* Δ*alc* mutant exhibited limited impairments in growth under ambient CO_2 and had a similar number of carboxysomes as WT cells ($p = 0.372$), albeit slightly enlarged ($p = 0.029$) (Figure 3.11, Table 3.3). Since Δ*rca* mutant lines in plants exhibit a high-carbon requirement for growth[191], a detailed ultrastructural analysis of WT and Δ*alc* strains of *F. diplosiphon* under air, with or without 3% CO_2 enrichment was performed, using TEM. Due to the moderately high light level of this chamber for *F. diplosiphon* growth, cells of both strains exhibited an irregular distribution of PL without CO_2 enrichment;

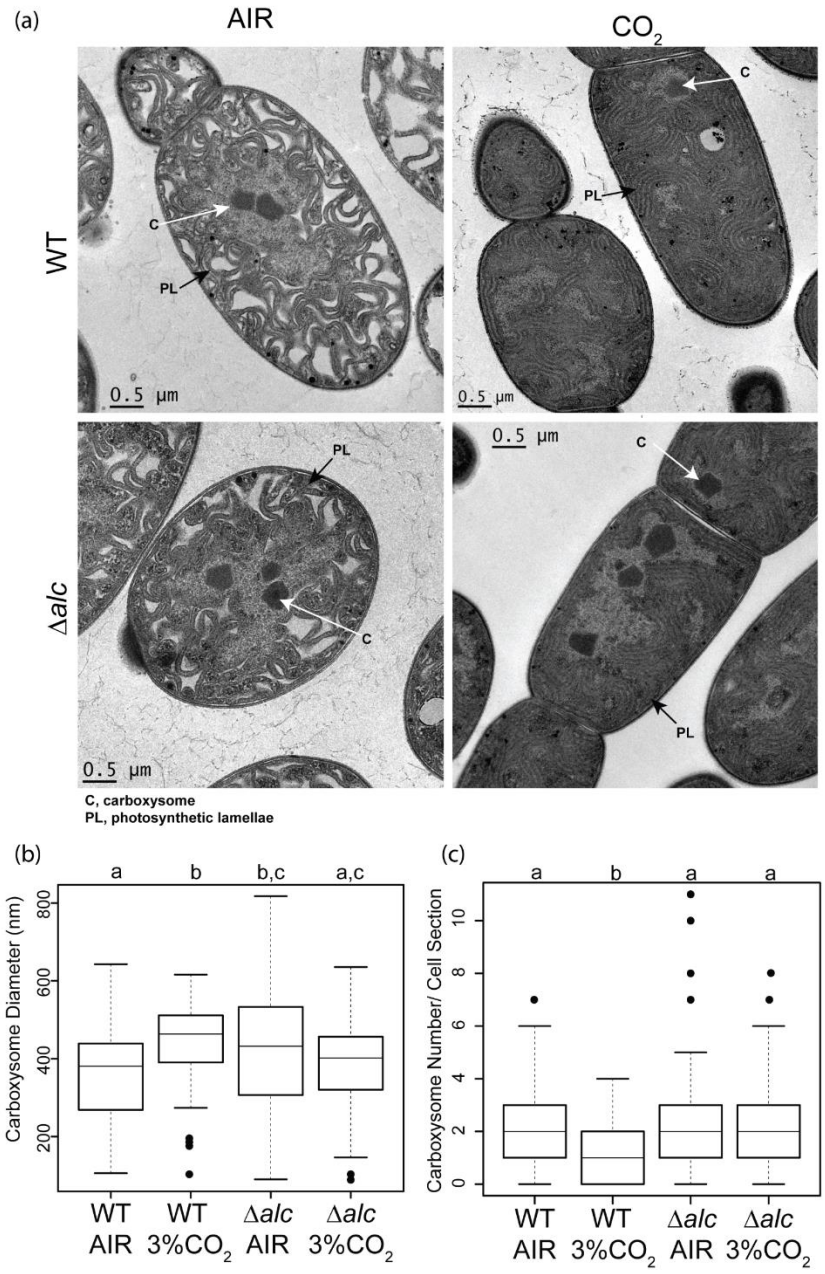


Figure 3.11: Δ*alc* phenotype in *F. diplosiphon*. (A) Representative images of TEM analysis of cellular ultrastructure of *F. diplosiphon* WT and Δ*alc* mutant strain. Cells were grown under white light with ambient air with or without the addition of 3% CO_2. C, carboxysomes (indicated by white arrows); PL, photosynthetic lamellae (indicated by black arrows). Bars, 0.5 µm. Carboxysome **(B)** size and **(C)** number measurements of WT and Δ*alc* strains under air and 3% CO_2. To determine size, the maximum diameters of carboxysomes from 20 cell sections of each strain were measured under each growth condition and are presented as a box plot. For carboxysome number, carboxysomes per cell section were counted in an additional 40 cell sections and are also presented as box plots. Carboxysomes per cell section is used as a proxy measurement for the number of carboxysomes in the whole cell and will be less than the total number per cell since it analyzes only a cross-section. Box plots were used as described in Rohnke et al.[170] to highlight the population dynamics of the measurements. Corresponding averages (± SE) can be found in Table 3.3. Statistical analyses were conducted using a Welch two-sample *t* test performed in R. Homogenous mean groups ($p > 0.05$) are indicated by identical letters above bars; different letters indicate a statistically significant difference ($p < 0.05$) from others.

Table 3.3: Quantification of average carboxysome sizes and average numbers of carboxysomes per cell section in WT and Δ*alc F. diplosiphon* from Figure 3.11.

Parameter	Value(s) for indicated strain			
	Air[a]		3% Enriched CO_2	
	WT	Δ*alc*[c]	WT	Δ*alc*
Carboxysome size (nm)[b]	362 ± 15	413 ± 17*	436 ± 19**	383 ± 16*
No. of carboxysomes/cell section	2.1 ± 0.2	2.5 ± 0.4	1.4 ± 0.1**	2.2 ± 0.2*
Sample size (n) for carboxysome size measurements	66	77	42	58
Sample size (n) for measurements of no. of carboxysomes/cell section	60	58	60	60

[a] Indicates conditions under which WT and Δ*alc* cells are grown, Air, ambient air; 3% Enriched CO_2, ambient air enriched with 3% CO_2

[b] Numbers for carboxysome size and carboxysome/cell section are represented as average ± SE.

[c] Statistical analyses, $p < 0.05$ indicated as follows: *, WT vs Δ*alc* in same condition; **, CO_2 vs. Air in same strain.

but a regular distribution of PL around the cell perimeter was generally recovered under 3% CO_2 enrichment (Figure 3.11A), although this recovery was not always observed in the Δ*alc* strain (Figure 3.12). Similar to *Syn*7942[199], the WT strain of *F. diplosiphon* showed a significant decrease ($p = 0.012$) in the number of carboxysomes per cell section when grown under CO_2 enrichment (Figure 3.12A & 3.12C). Additionally, there was a statistically significant increase ($p = 0.003$) in the carboxysomal diameter in WT under CO_2 enrichment conditions (Figure 3.12A & 3.12B). In contrast, the Δ*alc* mutant strain showed no significant differences in carboxysomes when comparing between conditions with or without CO_2 enrichment (number $p = 0.447$, size $p = 0.216$); instead carboxysome size and number per cell section under CO_2 enrichment were statistically indistinguishable from the WT strain grown under air (Figure 3.12B & 3.12C, Table 3.3).

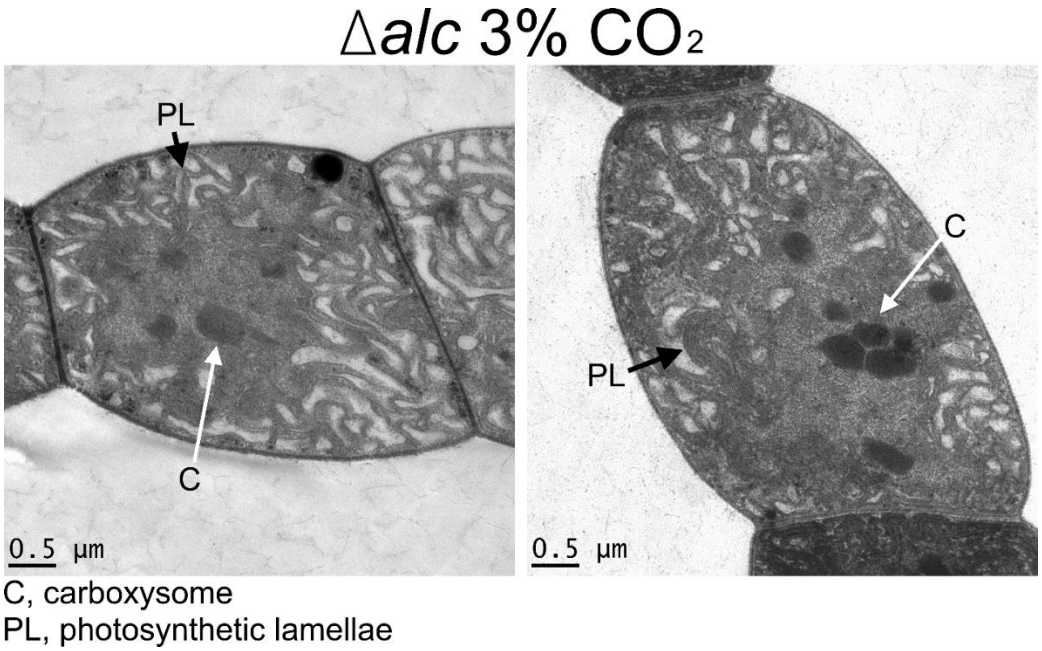


Figure 3.12: Diversity of cellular morphology in the Δ*alc* mutant strain when grown under air enriched for CO_2. Non-representative (i.e., occasionally occurring) images of TEM analysis of *F. diplosiphon* Δ*alc* mutant strain cellular morphology. Cells were grown under white light, with 3% CO_2. C, carboxysomes (indicated by white arrows); PL, photosynthetic lamellae (indicated by black arrows). Bars, 0.5 µm.

3.3 Discussion

Herein, a family of rubisco activase-like homologs exclusive to beta-cyanobacteria is characterized, which have been named activase-like cyanobacterial proteins (ALCs). It was shown that ALCs are broadly distributed across cyanobacteria from different taxonomic groups (Figure 3.1), absent only from the α-cyanobacterial C1 subclade, which contains Form IA rubisco and α-carboxysomes, and from the B2 subclade dominated by the unicellular *Chroococcales*[185]. It was shown that the SSLD fused to the C-terminus of nearly all ALCs is extremely similar to the SSLDs of the carboxysome protein CcmM (Figure 3.1C), and that ALCs retain key structural features of plant Rca (Figure 3.3 & Figure 3.5). Using recombinant *Fd*ALC, ATPase activity was confirmed *in vitro* (Figure 3.7B), and its localization to carboxysomes and close proximity to RbcL was demonstrated (Figure 3.6), revealing that it is required for normal carboxysome assembly under certain growth conditions (Figure 3.11). The present study also

reveals some intriguing biochemical behavior suggesting that the ALC may induce rubisco network formation, similar to activity recently described for CcmM (Figure 3.7D compared to Wang et al.[213]). Nevertheless, no effect upon rubisco activity was observed, and thus it remains a mystery whether it functions as an activase, but with a distinctive role in rubisco maintenance.

The presence of Rca-like proteins in cyanobacteria is puzzling because cyanobacterial rubisco had previously been shown to lack inactivation by RuBP[5,102,147,163], and thus would not need an activase. However, for the first time, RuBP sensitivity of a cyanobacterial rubisco was clearly demonstrated (Figure 3.7B). The concentration required to reach 50% of maximum inhibition was calculated to be 0.8 mM, which is about 10-fold higher than the reported cyanobacterial rubisco $K_{m(RuBP)}$[9,147,162]. In comparison, the concentration required to obtain 50% inhibition of the initial activation rate of *Rhodospirillum rubrum* rubisco ranged from 65 and 270 μM when the incubation performed at 2°C, and the lowest tested RuBP concentration of 0.4 μM resulted in more than 50% inhibition of spinach rubisco[83]. Hence, the cyanobacterial rubisco seems to be slightly less sensitive to RuBP inhibition than the Form II bacterial rubisco, and at least 1000-fold less sensitive than the plant rubisco. If such inhibition occurs *in vivo*, it would set the stage for requiring an activase to restore activity. The failure herein to detect neither rubisco activity enhancement by the ALC nor restoration of activity following RuBP-inhibition (Figure 3.7F) suggests that either the ALC has some other function distinct from that known for plant Rca, or that its activase function requires an accessory component or regulatory modification that was missing from our system. At the very least, the conservation of A144 (tobacco Rca numbering; *Fd*ALC A71), K247 (*Fd*ALC K175), and non-canonical pore-loop regions, which are characteristic of Rca activity[196] suggests a strong evolutionary connection to Rca function.

In addition to the conserved activase residues mentioned above, the ALCs show conservation of residues required for ATP hydrolysis (Figure 3.5E). Accordingly, *Fd*ALC showed ATPase activity *in vitro* (Figure 3.7B), similar to plant Rca, with a $K_{m(RuBP)}$ comparable to that of the plant enzyme (compare Figure 3.7B to report by Robinson and Portis[167]). The [ATP] measured in cyanobacterial cells[25,26], ranging from 165-230 nmol mg chl^{-1}, are about 10-fold higher than those measured in chloroplasts from spinach (25-40 nmol mg chl^{-1})[75,96]. However, there is currently no information about the ATP concentration inside the carboxysome, where the ALC is expected to function, moreover about how small metabolites such as ATP traverse the shell in any BMC[192,193].

The detection of FRET between the ALC-mT2 and the YFP-RbcL fusion proteins upon co-expression in *F. diplosiphon* (Figure 3.6C) implies that ALC and rubisco are located within 10 nm of each other[87], although this does not prove a direct interaction. As the interior of carboxysomes is expected to be densely packed with rubisco, which measures about 10 nm per L_8S_8 complex[22], and the average diameter of *F. diplosiphon* carboxysomes is about 360 nm, as measured from TEM (Figure 3.11, Table 3.3), it is very likely that the ALC is inside the carboxysome, in close proximity to rubisco. β-carboxysomes are synthesized from the inside out[38], with the core forming first by aggregation of rubisco via the M35 form of CcmM consisting of three or more tandem SSLDs, followed by recruitment of the icosahedral shell. It is hypothesized that the ALC gets targeted to the carboxysome through a similar interaction involving its SSLD.

Recently, it was demonstrated that M35 can induce aggregation of rubisco *in vitro*, observable as increased turbidity and evidenced by cryo-EM as the binding of one of the three SSLDs from M35 to a peripheral interface of rubisco involving both RbcL and RbcS[213]. Because the increase of turbidity could only be induced if a protein with two or more SSLDs was included in the aggregation assay, the authors suggest that rubisco network formation occurs with at least two of the three SSLDs in each M35 polypeptide binding to a different L_8S_8 rubisco complex, serving as a crosslinker. Surprisingly, it was demonstrated herein that *Fd*ALC addition to *F. diplosiphon* rubisco can induce the same phenomenon of increased turbidity (Figure 3.7D). Because the ALC SSLD shares an extremely high degree of sequence conservation with the SSLDs from CcmM (Figure 3.1C & Figure 3.2), it was expected to bind rubisco L_8S_8 peripherally in the same manner. Because the ALC is likely to function as a hexamer, SSLDs from the six monomers could link different molecules of rubisco L_8S_8 complexes, in order for ALC polypeptides to serve as crosslinks. Consistent with this hypothesis, the addition of genetic truncations of ALC consisting of both the SSLD and the AAA+ domain, without the linker between them, did not induce increase in turbidity (Figure 3.9); the effect only occurs with full-length ALC. Because ALC-mediated increased turbidity depends on the absence of ATP (Figure 3.7E), it is possible that increased rubisco networking in the carboxysome at night, while the cellular ATP concentration is low and rubisco is not active, increases rubisco stability. This mechanism of protection might be the cyanobacterial parallel to the plant rubisco protection by binding of the nocturnal inhibitor 2-carboxy-D-arabinitol 1-phosphate (CA1P), and other tightly bound inhibitors which protects the molecule from proteolysis and, perhaps, from reactive oxygen species[93,145]. Alternatively, or additionally if carboxysome integrity is dynamic, with the degree of rubisco packaging influenced by CO_2 levels[200], the ALC may play a role in rubisco

networking in the context of carboxysome assembly dynamics in response to CO_2 acclimation. This is consistent with our expression and ultrastructural data and will require detailed follow up studies. The sensitivity to ATP is a distinguishing trait of the ALC-mediated rubisco aggregation compared to the effect of M35, and might suggest different regulation of rubisco aggregation by the two, apparently redundant, systems.

Although there is one study of rubisco and its ALC in *Anabaena* spp., most studies of cyanobacterial rubisco involved unicellular species which lack an ALC such as *Synechocystis* sp. PCC 6714, and *Synechococcus* sp. PCC 7002[205], which could contribute to the reported lack of cyanobacterial rubisco inhibition. When an ALC gene was characterized in *Anabaena* sp. strain CA, several different cyanobacterial strains were screened for its presence[107]. Only heterocyst-forming cyanobacteria were found to contain an ALC in that study[107]. With the increasing availability of cyanobacterial genomic sequences it is now clear that many groups of cyanobacteria contain an ALC (this study and Zarzycki et al.[222]). It is enigmatic that ALCs are absent from the genomes of a large group of unicellular cyanobacteria (mostly subsection I[165] clustered in Branch B2[185]). Because the organisms in this branch do not share a single habitat or environmental milieu, it is reasonable to assume that the last common ancestor to this group has lost its *alc* gene, and this group has evolved without an ALC. This assumption is also supported by the ability to generate Δ*alc* mutants without severe growth defects (this study and Li et al.[103]). The correlation between the phylogenies of RbcL sequences and ALCs (Figure 3.1B) and the relationship of RbcS to the SSLD from the ALC suggest an ancient co-evolution of the two proteins. Furthermore, the nesting of CcmM-derived SSLD sequences within the clade of ALC-3-derived SSLDs, and the closer proximity of ALC-6 to true RbcS sequences (Figure 3.1C)

suggests that the SSLD from ALC may be the more ancient form, predating β-carboxysomes, with CcmM recruiting the SSLD from an ALC. Thus, ALCs may now be "vestigial", explaining their absence from some cyanobacteria.

This study establishes for the first time that the ALC is localized to the carboxysomes and is in close proximity to rubisco, which are encapsulated within the carboxysome microcompartment (Figure 3.6). This resembles the green alga *Chlamydomonas reinhardtii*, another unicellular aquatic chlorophotoautotroph which uses a functionally analogous microcompartment, the pyrenoid, for CO_2 concentration around rubisco[126], inside which Rca has been shown to co-localize[119]. Unlike higher plants, where an *rca* deletion causes severe growth impairment, deletion of *rca* in *C. reinhardtii* caused only a moderate decrease of growth rate under ambient CO_2, prompting a suggestion that the presence of a CCM may partially compensate for the loss of Rca[151]. Deletion of the *alc* gene in our study, as well as in *Anabaena variabilis*[103], similarly did not result in an HCR phenotype, supporting the interpretation that compartmentalization of rubisco may compensate or prevent the catalytic inefficiencies of rubisco. Upon C_i-upshift from air to high CO_2, the carboxysomes of *F. diplosiphon* Δ*alc* mutant failed to decrease in amount per cell nor showed an increased size, in contrast to WT[199] (Figure 3.11B & 3.11C), nor did *ccmM* transcript abundance increase (Figure 3.10). Together with the observation that *alc* transcript abundance was increased in response to higher CO_2 levels (Figure 3.10), these observations reinforce the idea that ALCs play an important role in metabolic response to carbon (and possibly ATP) availability by influencing rubisco packaging within the carboxysome. It is proposed that while the Δ*alc* mutant strain is able to compensate for greater carbon stress under standard conditions, the protein is required for acclimation to help mitigate carbon stress. This

culminates in a distinct lack of response to the CO_2 surplus. Such carbon stress would be expected if the rubisco in these cells was less active than in WT. The upregulation of *alc* in response to elevated CO_2 further supports ALC's role in utilizing a CO_2 surplus.

The ALC does not appear to function as a canonical rubisco activase, enzymatically acting on rubisco. Instead it influences the metabolic response to CO_2 availability at the level of the carboxysome. Notably, the seemingly subtle benefits that could be afforded by the ALC strengthen the view that this protein serves as an evolutionary link between the fast, promiscuous carboxysomal rubisco and the Rca-requiring rubisco of higher-order plants. Understanding the involvement of the ALC in regulation of photosynthesis through its modulation of rubisco networking and carboxysome acclimation to CO_2 levels could contribute to efforts to improve productivity in cyanobacteria, and potentially in plants.

3.4 Materials and methods

3.4.1 Bioinformatic analyses

353 cyanobacterial genomes retrieved from the Integrated Microbial Genomes database (IMG; https://img.jgi.doe.gov/) were profiled for the presence of the AAA+ domain (pfam00004). The resulting 4562 sequences were subsequently scored for matches to a position-specific substitution matrix (PSSM) generated for each of three conserved pore-loop regions identified in a previously-published alignment of 22 ALCs[222]. These regions of the conserved pore-loops were found to provide the best distinction between Rca/ALC and other unrelated AAA+ proteins. PSSM generation and scoring was performed using BioBike (http://biobike.csbc.vcu.edu:8003/biologin). Pore-loop sequences used for the generation of

PSSM – Loop 1 – pos. 61 – 88 (*Fd*ALC position numbering); Loop 2 – pos. 98 – 142; Loop 3 – pos. 150 – 179. Sequences which contained all 3 loops were identified as ALCs. These sequences were aligned with several previously investigated reference Rca sequences from plants and algae[196]. The sequences of 133 identified ALCs and 10 green-type Rca proteins (Table S1) were aligned using MUSCLE[47], and resulting multiple sequence alignments (MSA) were trimmed to retain only the AAA^+ domain and edited manually. Phylogenetic analysis of aligned AAA^+ domain was done using PHYML[69] with empirically-determined substitution rates, SH-like branch supports, and the automated method for tree improvement. Sequence similarity groups resembling subtype clades were identified manually by long internal branch lengths (ALC-1 to ALC-6); eight outliers were not assigned to any clade.

The MSA of 133 ALCs was used to generate conservation scores mapped to the homology model using the ConSurf server (http://consurf.tau.ac.il/2016/). ConSurf models were rendered in PyMol.

Full length RbcL sequences were retrieved from IMG for the 353 cyanobacterial genomes using BLAST queried with the *F. diplosiphon* RbcL. Rubisco-like protein sequences were removed based on shorter length and annotation as 2,3-diketo-5-methylthiopentyl-1-phosphate enolase. The resulting 335 RbcL sequences were aligned with 13 green-type reference RbcLs (Table S1) using MAFFT[88] with iterative refinement based on local pairwise alignment (L-INS-i), and a phylogenetic tree was constructed using RaxML[195] with the rapid hillclimbing search and an automatic amino acid substitution model. Phylogenetic trees were visualized using Archaeopteryx[226].

For sequence analysis of RbcS and SSLD, proteins with pfam00101, corresponding to RbcS and SSLD were retrieved for 327 cyanobacteria genomes in IMG. A total of 107 representatives of each RbcS and CcmM, and 131 ALC sequences were used as an initial target for HMMSearch[153]. An HHM generated from 189 representative sequences containing pfam00101 in the RP15 database (http://pfam.xfam.org/family/PF00101), using HMMBuild was used as the HMMSearch query. The SSLDs were retrieved from CcmM and ALC sequences based on their match to the HMM, and representative sequences were aligned using MUSCLE[47]. The phylogenetic tree was constructed in PHYML[69]. HMM sequence logos were generated using Skylign.org[218] using RbcS sequences from 223 β cyanobacteria, 111 CcmM-SSLD domains from 32 β-cyanobacteria and 108 ALC-SSLD sequences.

3.4.2 Protein homology modeling

For ALC, the RaptorX (http://raptorx.uchicago.edu/) web server was used to generate a multi-domain homology model based on both Rca from tobacco (PDB: 3T15) and RbcS from *T. elongatus* (PDB: 2YBV chain B) with a flexible, linker region in between. Further homology models were made using Swiss-Model[216] (https://swissmodel.expasy.org/) based on the SSLD of CcmM (PDB: 6MR1) for the final SSLD model in *Fd*ALC (residues 334 – 424), as well as based on the hexameric tobacco Rca structure (PDB: 3ZW6) for the AAA+ domain of *Fd*ALC (residues 1 – 288) and RbcS from *Syn*6301 (PDB: 1RBL chain M) for *Fd*ALC residues 317 – 424. In addition, alignment scores between two sequences were calculated using the LAlign webserver (https://embnet.vital-it.ch/software/LALIGN_form.html) in order to evaluate candidate template structures and to compare primary structure conservation. PDBsum

(https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html) was used to generate schematics of protein secondary structure.

3.4.3 Cloning and growth conditions

Standard molecular biology techniques were used for molecular cloning. Detailed lists of strains, plasmids and oligonucleotides used in this study are provided in Lechno-Yossef et al.[101].

E. coli strains were grown in LB supplemented with appropriate antibiotics at standard concentrations[8]. Strains DH5α or DH5αMCR were used for plasmid constructions. DH5αMCR was used to harbor methylating, conjugative and cargo plasmids for conjugation with *F. diplosiphon*[50]. Standard antibiotic concentrations of ampicillin at 100 µg mL^{-1}, chloramphenicol at 25 µg mL^{-1}, gentamycin at 50 µg mL^{-1}, kanamycin at 50 µg mL^{-1}, or spectinomycin at 100 µg mL^{-1} were used.

SF33[42] was used for strain constructions, and is treated as the WT comparison in all *in vivo* experiments. *F. diplosiphon* strains were grown in BG11/HEPES[165] at 28 - 30°C in the light (ca. 25-30 µmol photons m^{-2} s^{-1}) and enriched with 3% CO_2, unless otherwise stated, in shaken liquid cultures. For selection of mutants on plates, BG11/HEPES was solidified with 1.2% Difco agar, and 3 g of sodium thiosulfate per liter was added to the medium. Antibiotic concentrations used for selection of *F. diplosiphon* mutants were kanamycin at 25-50 µg mL^{-1}, or spectinomycin at 10 µg mL^{-1}.

3.4.4 Construction of a methylating plasmid

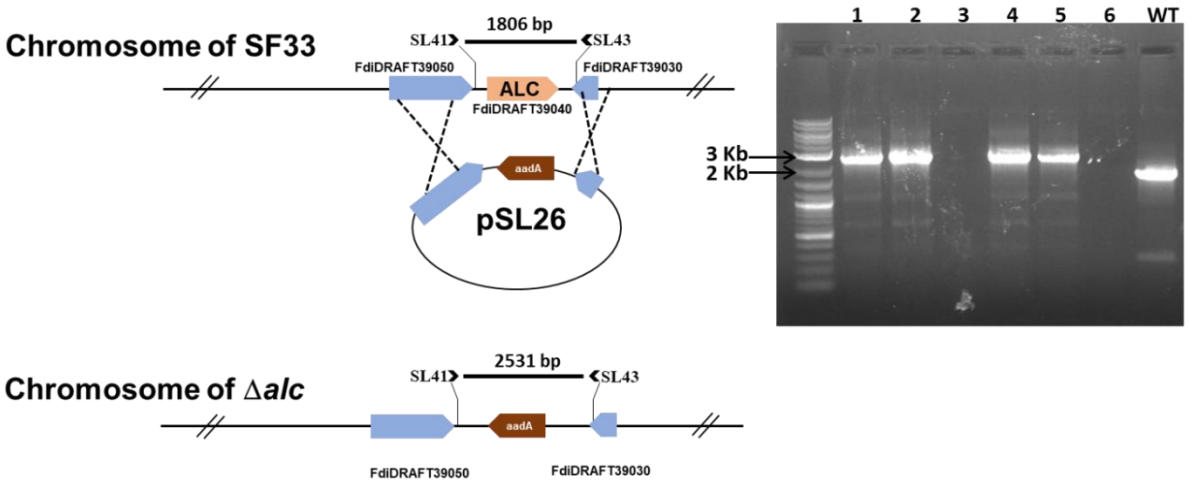
F. diplosiphon contains three known restriction enzymes – *Fdi*I is an isoschisomer of *Ava*II; *Fdi*II - isoschisomer of *Fsp*I and *Fdi*III - isoschisomer of *Sph*I. The current methylating plasmid used for transformation - pJCF173, carries methylases that protect only *Fdi*I and *Fdi*II recognition sequences[43]. To include a third methylase, protecting the *Fdi*III recognition sequence, a new methylating construct was made. The plasmid pRL518[51], containing M.Eco47II, was used as the basis for this construct. Constructs are described in Lechno-Yossef et al.[101]. Briefly, the chloramphenicol resistance marker in pRL518 was replaced by a gentamycin resistance cassette to allow replication and selection along with cargo plasmids containing chloramphenicol-resistance markers. Genes for M.*Fdi*II (fdiDRAFT68880) and M.*Fdi*III (fdiDRAFT17080) were PCR-amplified and cloned downstream of the *Amaranthus hybridus psbA* promoter, a promoter that has been shown to be constitutively expressed in *E. coli*[49] (Elhai, 1993). The genes with the promoter were then cloned in gentamycin-resistant version of pRL518, and the new construct – pSL17 was tested for protection against digestion with *Fsp*I and with *Sph*I.

3.4.5 Construction of mutant strains and strains expressing fluorescent fusion proteins

All strains were generated by di-parental mating between an *E. coli* DH5αMCR, containing the conjugative plasmid pRL443, methylating plasmid, pSL17 and a desired cargo plasmid as described[50]. Fusion fluorescent proteins were expressed from the GL-inducible promoter for phycoerythrin-coding gene *cpeE,* or the core phycobilisome gene promoter *apcA* on a replicating plasmid based on the native plasmid from *F. diplosiphon*, or an RSF1010-based compatible plasmid. To express the ALC-mT2 fusion, pSN7 was generated in which the coding sequence

was expressed from the GL-inducible promoter on a plasmid derived from pPL2.7[175]. The plasmid pSL198 is based on an RSF1010 replicon and contains the YFP-RbcL coding sequence from the *apcA* promoter. Detailed plasmid construction is presented in Lechno-Yossef et al.[101].

The *alc* knockout-mutant strain was generated by homologous recombination of pSL26, selected on spectinomycin. Complete segregation of the double recombinants was tested by PCR (Figure 3.12).



.

Figure 3.13: Construction of Δ*alc* mutation by homologous recombination and verification of complete segregation. A suicide plasmid carrying a replacement of the alc gene by the resistance gene to streptomycin and spectinomycin was used to replace the gene in the chromosome of F. diplosiphon. Colony PCR was performed on 6 bacteria-free spectinomycin-resistant colonies using primers SL41 and SL43[101]. The expected PCR size for the fully segregated mutant is 2531 base pairs, and that of the WT is 1806 base pairs. Four out of the 6 tested colonies appear to be fully segregated deletion mutants of the ALC gene.

3.4.6 Confocal scanning laser microscopy

Cell cultures were induced by growth under green light (10 – 15 µmol photons m^{-2} s^{-1}) for 7 – 10 d and were immobilized on an agarose-covered microscope slide. Slides were observed using Olympus FLUOVIEW FV1000 confocal laser scanning microscope using differential interference contrast optics and fluorescence excitation and emission filters. A 60X, 1.42-

numeric-aperture oil immersion objective lens was used. For differential interference contrast imaging, the 488 nm laser was used. For detection of blue-channel fluorescence from the mT2, a 405 nm laser diode was used for excitation and the emission was detected with a band pass filter BA430-470. For detection of yellow-channel fluorescence from YFP, a 488 nm Argon gas laser was used for excitation and the emission was detected with a band pass filter BA535-565. FRET between *Fd*ALC-mT2 and YFP-RbcL was conducted as described by Karpova and McNally[87]. Briefly, an image of fluorescence in a certain field was taken in the blue and yellow channel. Then the yellow signal in a specific region of the imaged field was photobleached by the laser. The same field was imaged again. Olympus FV1000 software was used to calculate the change in fluorescence intensity in both channels in the bleached region as well as in two control regions in each visualized field. Averages from 10 different images are presented in Figure 3.7C. All images were visualized using ImageJ[176]. Pseudocolors were applied as follows: cyan for mT2 channel, yellow for YFP channel.

3.4.7 Protein expression and purification from *E. coli*

The full length *FdALC* was cloned in a pCDFDuet1 vector (Novagene, Darmstadt), with a C-terminal 6XHisTag[101], and the clone was denoted pAB2. Sequence of *F. diplosiphon* CcmM starting from amino acid 229 was cloned in the NdeI site of pET28a (Novagene, Darmstadt), generating pSL286 that carries N-terminally His-tagged M35. The *F. diplosiphon rbcL* gene was cloned in the first polylinker of pCDFDuet1 with a C-terminal StrepII tag. The two genes downstream in the genome, *rbcXS*, were then cloned immediately downstream with their native intergenic regions, generating pAB9. For expression of rubisco in *E. coli*, the *F. diplosiphon*

gene for the chaperone Raf1[57] was cloned in the arabinose inducible plasmid pBAD24[72], generating pAB24.

For expression of *Fd*ALC and M35, cultures of *E. coli* strain BL21 DE3 (Invitrogen, Carlsbad, CA, USA) harboring pAB2 or pSL286 were grown to optical density at 600 nm of 0.4-0.7 at 37°C, induced with 50 μM IPTG, and grown for additional 18 – 24 h at 25°C. For expression of rubisco, BL21 DE3 strain containing pAB9, pAB24 and pGro7 (containing GroES and GroEL[139] (TaKaRa Bio USA, Mountain View, CA). Upon inoculation, chaperones were induced with 0.2% (w/v) L-arabinose, and cultures were grown to optical density at 600 nm of 0.4-0.7 before induction with IPTG as above. Cultures were harvested by centrifugation and pellets were either processed immediately or stored at -20°C until purification. For His-tagged protein purification, cell pellets were suspended in lysis buffer (50 mM Tris, pH 8; 200 mM $MgSO_4$; 10% glycerol) containing EDTA-free SigmaFast protease inhibitor cocktail at the recommended concentration (Sigma, St. Louis, MO) and approximately 10 μg mL^{-1} DNAseI at 1 – 2 mL per g fresh cell weight. Cells were lysed by passing twice through a cell disruptor (Constant Systems, Aberdeenshire, UK) at 15 kPSI. Cell lysates were clarified by centrifugation at 45,000 g for 30 min, and the clear lysate was passed through a 0.45 μm filter, before adding imidazole to a final concentration of 100 mM and loading on a HisTrap column (GE Healthcare, Little Chalfont, UK), equilibrated with Buffer A (50 mM Tris, pH 8; 200 mM NaCl; 10% glycerol) containing 100 mM imidazole, attached to an AKTA Pure FPLC (GE Healthcare, Little Chalfont, UK). DTT at final concentration of 1 mM was added to all buffers used for *Fd*ALC purification, but not to M35 purification buffers. For *Fd*ALC purification, column was washed with 5 column volumes of buffer A containing 100 mM imidazole and 5 column volumes of buffer A

containing 175 mM imidazole, the sample was eluted in buffer A containing 500 mM imidazole. For M35 purification, after washing the column with 5 column volumes of buffer A containing 100 mM imidazole, the protein was eluted with a gradient of 100 mM to 500 mM imidazole in buffer A over 10 column volumes. No further purification was done. For purification of rubisco, cell pellets were suspended in rubisco purification buffer (50 mM phosphate buffer, pH 7.6; 200 mM KCl; 1 mM EDTA; 1 mM DTT) supplemented with SigmaFast protease inhibitor cocktail at the recommended concentration (Sigma, St. Louis, MO) and approximately 10 $\mu g\ mL^{-1}$ DNAseI at 1 – 2 mL per g fresh cell weight. Cell suspensions were passed twice through French Press at 1100 PSI. The lysate was clarified by centrifugation as above and loaded on StrepTrap HP, 5 mL (GE Healthcare, Little Chalfont, UK), equilibrated with rubisco purification buffer. The column was washed with 5 column volumes of rubisco purification buffer and eluted in the same buffer containing 5 mM desthiobiotin (Sigma, St. Louis, MO).

For results presented in Figure 3.8, rubisco was purified from *F. diplosiphon* cultures. *F. diplosiphon* cells harvested from 1 L dense culture, lysed in buffer containing 25 mM Bicine pH 7.6, 5 mM EDTA, 2 mM DTT, DNAseI and a protease inhibitor cocktail (Sigma, St. Louis, MO) by passing 3 times through French Press. The lysate was clarified by 1-hour centrifugation at 20,000 g. The supernatant was treated with 20% saturated ammonium sulfate, and precipitated proteins were discarded. Rubisco was precipitated by 50% saturation ammonium sulfate solution, suspended in lysis buffer treated with 0.5% triton X-100, and precipitated again with 20% PEG-6000. Precipitated proteins were then suspended in ion exchange loading buffer containing 100 mM phosphate and 1 mM DTT and loaded on a MonoQ HR 16/10 column (GE healthcare, Little Chalfont, UK), attached to an AKTA Pure FPLC (GE Healthcare, Little

Chalfont, UK). Gradient with elution buffer containing 100 mM phosphate, 1 M KCl and 1 mM DTT was used to elute rubisco.

Proteins were observed on SDS-PAGE gel stained with coomassie blue, and the concentration was quantified using the BCA assay kit (Sigma, St. Louis, MO), against bovine serum albumin standard.

3.4.8 ATPase activity assay

ATPase activity was performed by coupling ADP release to NADH oxidation via pyruvate kinase and lactate dehydrogenase, as described[202]. The reaction buffer contained 50 mM Tris-HCl pH 8.0, 20 mM KCl, 5 mM $MgCl_2$, varying concentrations of ATP, 1 mM phosphoenolpyruvate, 0.3 mM NADH, 12 U/mL pyruvate kinase, and 12 U/mL lactate dehydrogenase. The reaction was initiated by the addition of 2.5 μM recombinant ALC, in a final reaction volume of 0.5 mL. The change in absorbance was followed at 340 nm for 2 min using Agilent Cary 60 UV-Vis Spectrophotometer (Agilent, Santa Clara, CA). Activity was calculated assuming 1:1 ratio of ADP:NADH and a 6.22 $mM^{-1}cm^{-1}$ extinction coefficient.

3.4.9 Rubisco and activity assays

Rubisco activity was measured by rate of incorporation of $^{14}CO_2$. Reactions were carried out with purified rubisco at varied concentrations described below for individual experiments, and conducted at 25°C. To test the inhibition by RuBP (results presented in Figure 3.7C), rubisco (500 μg mL^{-1}) was first incubated for 10 minutes with varied concentrations of RuBP, and placed on ice until activation. The inhibited sample was then activated for 10 min at room temperature

with 20 mM $MgCl_2$, 20 mM $NaHCO_3$. The activity of the activated enzyme with or without pre-incubation with RuBP was measured by adding 20 μL of the activated enzyme into 80 μL assay buffer containing 50 mM EPPS pH 8, 5 mM $MgCl_2$, 0.2 mM EDTA, 1 mM RuBP and 15 mM $NaHCO_3$ and trace amount of $NaH^{14}CO_3$ (10 μC_i / mL). The reaction was stopped after 1 minute by adding 100 μL 1 M formic acid and drying on a hot plate. Each reaction was done in triplicates. Acid-stable radioactivity was measured by a scintillation counter model Tri-Carb 2800TR (Perkin Elmer, Waltham, MA). For testing activation by the ALC (Results presented in Figure 3.7F), rubisco (0.25 μM) that has been pre-incubated with or without 4 mM RuBP for 10 minutes, was incubated with 50 mM EPPS pH 8, 20 mM $MgCl_2$, 10 mM $NaHCO_3$, 40 U mL^{-1} creatine phosphokinase, 4 mM creatine phosphate, 2 mM DTT, and 10% PEG-3350. Five mM ATP was included in some samples, as specified in figure legends. Either BSA (as a negative control) or recombinant ALC were included as specified in Figure 3.7. After 10 min activation 20 μL of the activated rubisco were mixed with 80 μL of assay buffer, incubated for 2 min and stopped as described above.

3.4.10 Turbidity assays

Turbidity assays were conducted as described[213]. The assay was conducted in 100 μL volume in buffer containing 50 mM Tris, pH 8; 10 mM Magnesium acetate; 50 mM KCl; 5 mM DTT. Rubisco (0.25 μM, unless otherwise stated) was added to the assay buffer and placed in a spectrophotometric cuvette. Absorbance at 340 nm was followed using the kinetics setting on an Agilent Cary 60 UV-Vis Spectrophotometer (Agilent, Santa Clara, CA), and the additional proteins, *Fd*ALC or M35 at 2 μM final concentration were mixed in the cuvette after initiation of the measurement. Change of absorbance was followed for 10 min.

3.4.11 qPCR Analysis

F diplosiphon cells were grown with and without the enrichment of 3% CO_2 for 3 d and were normalized to an OD_{750} of 0.40, then cells grown under 3% CO_2 were split into two groups. One group (High C_i) was returned to grow under 3% CO_2 enrichment, while the other (C_i Stress) was moved to grow without CO_2 enrichment to simulate low C_i stress according to the methods of Wang et al.[214]). Cells grown without CO_2 enrichment (Air) were treated as the intermediate C_i samples. Cells were then grown for 19 hrs, cooled briefly with liquid N_2, then harvested by centrifugation at 5,125 x *g* at 4°C for 10 min. The cell pellet was transferred to a microfuge tube, centrifuged at 13,000 x *g* at 4°C for 5 min, decanted, and then flash frozen at -80°C. mRNA was extracted with the addition of Trizol reagent (1 mL), vortexing, incubation at 95°C for 8 min, incubation on ice for 8 min, addition of chloroform (200 µL) with brief vortexing, centrifugation at 13,000 x *g* for 15 min, and transfer of ~500 mL supernatant to a new tube. For RNA purification, an equal volume of isopropanol was added, and the sample was incubated at room temperature for 15 min, followed by centrifugation at 13,000 x *g* for 8 min. The supernatant was decanted, then 70% ethanol was added, and the sample was dried for 5-10 min following centrifugation at 13,000 x *g* for 8 min and decanting the supernatant. The RNA pellet was resuspended in 70 µL nuclease-free water and treated with TURBO DNA-free kit (Invitrogen, Madision, WI) according to manufacturer's directions. mRNA was quantified using a NanoDrop ND-1000 Spectrophotometer, diluted to a concentration of 100 ng/µL, and 250 ng was used for a 25 µL reverse transcription reaction using qScript cDNA SuperMix (Quantabio, Beverly, MA) kit according to the manufacturer's directions. No reverse transcription reactions were also performed without the addition of reverse transcriptase to evaluate genomic contamination.

qPCR was prepared using Fast SYBR-Green master mix (Applied Biosystems, Foster City, CA) according to the manufacturer's directions in 384-well plates (Applied biosystems). *alc*, *ccmM*, and the endogenous control *orf*10B were amplified using primers specified in Lechno-Yossef et al.[101]. Samples were run and analyzed using the ABI QuantStudio 7 Flex PCR system (Applied Biosystems) in fast mode. Samples were run in triplicate, with 5 biological replicates from two independent experiments. Statistics were run using a Welch one-sample *t* test performed in Excel to generate 95% CI.

3.4.12 TEM Analysis

TEM analysis and quantification of carboxysome size and number were performed according to the methods of Rohnke et al[170]. In total, ~60 cell sections were analyzed for carboxysome number in each strain under each condition, with carboxysome diameters measured in 20 of these cell sections.

3.5 Acknowledgements

This work was supported by the Office of Science of the US Department of Energy DE-FG02-91ER20021. The authors would like to thank Dr. Sarah E. Newnham and Dr. C. Raul Gonzalez-Esquer for help with fluorescent tag cloning, Dr. Shailendra P. Singh for assistance with structural modeling of the *F. diplosiphon* ALC, Dr. Dayana Muzziotti and Dr. Melinda Frame, for help with confocal microscopy, Dr. Sarathi Wijetilleke for help with rubisco assays, Mr. Damien Sheppard and Mr. Donovan Prince Smith for technical laboratory assistance, Dr.

Thomas D. Sharkey and Dr. Berkley Walker for assistance with ATP hydrolysis and rubisco assay experimental design, interpretation of results and critical reading of the manuscript.

3.6 Author contribution

S.L. and B.A.R designed and conducted the research, analyzed and interpreted data, and wrote the article. A.C.O.B conducted biochemical assays, M.R.M assisted in bioinformatics analyses and wrote the article, C.A.K and B.L.M. designed the research, analyzed and interpreted the data, and wrote the article.

CHAPTER 4

Binding Options for the Small Subunit-Like Domain of Cyanobacteria to Rubisco

This chapter contains information submitted for publication in:

Rohnke, B. A., Kerfeld, C. A., and Montgomery, B. L. (2019) Binding options for the small subunit-like domain of cyanobacteria to rubisco. *Front. Microbiol. Under Review*.

modified to incorporate the supplemental information into the body of the text, renumber figures, tables, and references to be consistent with the dissertation, and use abbreviations defined in the KEY TO ABBREVIATIONS.

4.1 Abstract

Two proteins found in cyanobacteria contain a C-terminal domain with homology to RbcS. These SSLDs are important features of CcmM, a protein involved in the biogenesis of carboxysomes found in all β-cyanobacteria, and a rubisco activase homologue (ALC) found in over a third of sequenced cyanobacterial genomes. Interaction with rubisco is crucial to the function of CcmM and is believed to be important to ALC as well. In both cases, the SSLD aggregates rubisco, and this nucleation event may be important in regulating rubisco assembly and activity. Recently, two independent studies supported the conclusion that the SSLD of CcmM binds equatorially to L_8S_8 holoenzymes of rubisco rather than by displacing an RbcS, as SSLD structural homology would suggest. We use sequence analysis and homology modeling to examine whether the SSLD from the ALC could bind the large subunit of rubisco either via an equatorial interaction or in an RbcS site, if available. We suggest that the SSLD from the ALC of *Fremyella diplosiphon* could bind either in a vacant RbcS site or equatorially. Our homology modeling takes into account N-terminal residues not represented in available crystal structures that potentially contribute to the interface between RbcL and RbcS. Here, we suggest the perspective that binding site variability as a means of regulation is plausible and that the dynamic interaction between the RbcL, RbcS, and SSLDs may be a means of carboxysome assembly and function.

4.2 Introduction

Due to the evolution of a CCM, cyanobacteria are able to significantly contribute to global carbon fixation, despite the comparatively low atmospheric CO_2 levels relative to their first appearance on Earth some 3.5 billion years ago[177,220]. The CCM serves to significantly increase

the flux of inorganic carbon into proteinaceous bacterial microcompartments called carboxysomes. Carboxysomes serve to encapsulate rubisco and the shell acts as a semi-permeable barrier to CO_2 escape, allowing rubisco to function under high substrate levels[46]. In the case of β-carboxysomes, which are present in cyanobacteria that express Form IB rubisco, synthesis occurs from the inside-out beginning with condensation of rubisco and carboxysomal protein CcmM into a liquid matrix[38,137,213].

The structure of CcmM is key to its nucleation of carboxysomal cargo. The C-terminus of CcmM contains 3-5 repeats of a domain that is homologous (around 60-70% similarity) to RbcS—denoted a an SSLD[116,156]. The SSLD repeats domain-containing portion of CcmM can also be independently transcribed through an internal ribosome entry site. Both CcmM forms—full-length M58 and truncated M35—are necessary for normal carboxysome biogenesis[113,114].

SSLDs were implicated in the interaction between rubisco and CcmM[44,112] and were long hypothesized to bind in place of RbcS in rubisco complexes[54,113,160]. However, recent structural work on SSLDs demonstrates equatorial binding of CcmM to L_8S_8 rubisco holoenzymes[171,213]. This binding appears to be driven largely by electrostatic interactions and affinity for CcmM is not affected even when RbcS binding is partially compromised[171].

SSLDs also appear as C-terminal domains in rubisco activase homologues (ALC) found in many cyanobacteria[222]. Recently, the ALC was shown to localize proximal to rubisco in the carboxysome and induce rubisco aggregation, much like M35[101]. Together, these findings

provide strong evidence that the SSLD of ALC binds to rubisco and can induce liquid-liquid phase separation.

As recent findings indicate that SSLDs do not displace RbcS in L_8S_8 rubisco holoenzymes and instead bind equatorially, it is puzzling why there would be conservation of the RbcS-like secondary and tertiary structure that facilitates interactions with RbcL. Some of the conserved residues from RbcS may fill repurposed roles in the SSLD-unique equatorial binding position, thus driving conservation of these features. Others, though, suggest that RbcS displacement may be possible. We decided to systematically homology model the SSLD and RbcL and compare the interfaces in order to evaluate the plausibility of equatorial versus RbcS substitution as a mode of binding[171,213]. We modeled the SSLD found in the ALC of *F. diplosiphon* (*Fd*ALC SSLD) and analyzed the number of predicted interactions and free energy of solvation when the SSLD binds at the RbcS site (i.e., binds an empty site or displaces RbcS) or binds equatorially. We suggest that while equatorial binding was favored for CcmM in *Syn*7942 which lacks an ALC homolog, the *Fd*ALC SSLD had similar interface features in both positions. We propose that the *Fd*ALC could bind either equatorially or in place of RbcS and suggest that the current models of equatorial SSLD may be a part of a larger set of possibilities depending on specific proteins, for example whether or not the cyanobacterium contains an ALC, and perhaps is reflective of the recently uncovered diversity of cyanobacterial RbcL subunits[101].

4.3 Methods

4.3.1 Protein homology modeling

The structures of the *F. diplosiphon* proteins were generated by homology modeling. For *Fd*ALC SSLD, the Swiss Model web server (https://swissmodel.expasy.org)[20,216] was used to generate a model for amino acid residues 317 – 424 based on *Syn*7942 CcmM SSLD1 in the reduced (PDB: 6HBB) form as well as *Syn*6301 RbcS (PDB: 1RBL, Chain M). Additionally, a homology model of *F. diplosiphon* L_8S_8 rubisco was made using *Syn*6301 rubisco (PDB: 1RBL) as a template. Alignment scores between two sequences were calculated using the LAlign webserver (https://embnet.vital-it.ch/software/LALIGN_form.html) in order to evaluate candidate template structures and to compare primary structure conservation.

4.3.2 MSA of ALC SSLDs

The MSA of 141 ALCs from cyanobacteria described in Lechno-Yossef et al.[101] was trimmed to remove the ATPase domain, then the remaining regions (linker and SSLD) were re-aligned with a low gap cost at the end of sequences in CLC Sequence Viewer. This allowed for alignment of the SSLD region despite significant variations in the sizes of linkers between species, which were then trimmed to match the SSLD region identified in *Fd*ALC (residues 317 – 424, corresponding to residues 1 – 107 of the SSLD). An MSA was also generated for RbcS for each of the 128 organisms that had both a full length SSLD and an annotated RbcS sequence. MSA for RbcS and SSLD were visualized and compared using HMM logos[178] generated on Skylign (http://skylign.org/).

4.3.3 Analysis of protein-protein interactions

Using the homology model for L_8S_8 *F. diplosiphon* rubisco, the two *Fd*ALC SSLD models were aligned to $RbcS_1$ in PyMol, and structures were generated containing each SSLD replacing $RbcS_1$. Another structure aligned *Syn*7942 CcmM SSLD1 in complex with rubisco (PDB: 6HBC) to the *F. diplosiphon* rubisco, and then *Fd*ALC SSLD (reduced) was aligned to the CcmM SSLD resulting in a *F. diplosiphon* rubisco model with *Fd*ALC SSLD in the M position (Fig 4.1A). The *Syn*7942 CcmM SSLD1 structure was also used to replace the $RbcS_1$ position in the *Syn*6301 rubisco L_8S_8 structure. Local refinement of structures was performed using Rosetta 3.4. Structures were subjected to the docking prepack protocol followed by the generation of 1000 decoys using the docking protocol in docking local refine mode with the SSLD as the mobile target[41,68,212]. Based on interface score, the top 200 structures were clustered by pairwise RMSD with a 1 Å cutoff using energy-based clustering in Rosetta 3.4[77]. In all cases, the structure with the lowest interface score belonged to the largest cluster and was selected for use in downstream analysis.

These structures, as well as the *F. diplosiphon* rubisco model and PDB: 6HBC, were analyzed using the Profunc web server (https://www.ebi.ac.uk/thornton-srv/databases/profunc/)[100]. Interactions involving the RbcS and SSLD were compared for each structure. Further analyses were performed using the Pisa webserver (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver)[98] to calculate the solvation free energy gain (ΔG) upon formation of the interfaces for each structure.

4.4 Binding at the RbcS$_1$ position

Each RbcS in an L_8S_8 assembly forms four unique protein-protein interfaces, three with the surrounding RbcL subunits and a fourth with a proximal RbcS (Figure 4.1A & 4.1B). The number of interfaces and predicted residue-level interactions were comparable to results from molecular dynamic simulations using *C. reinhardtii* rubisco[117] and crystallographic structures[95].

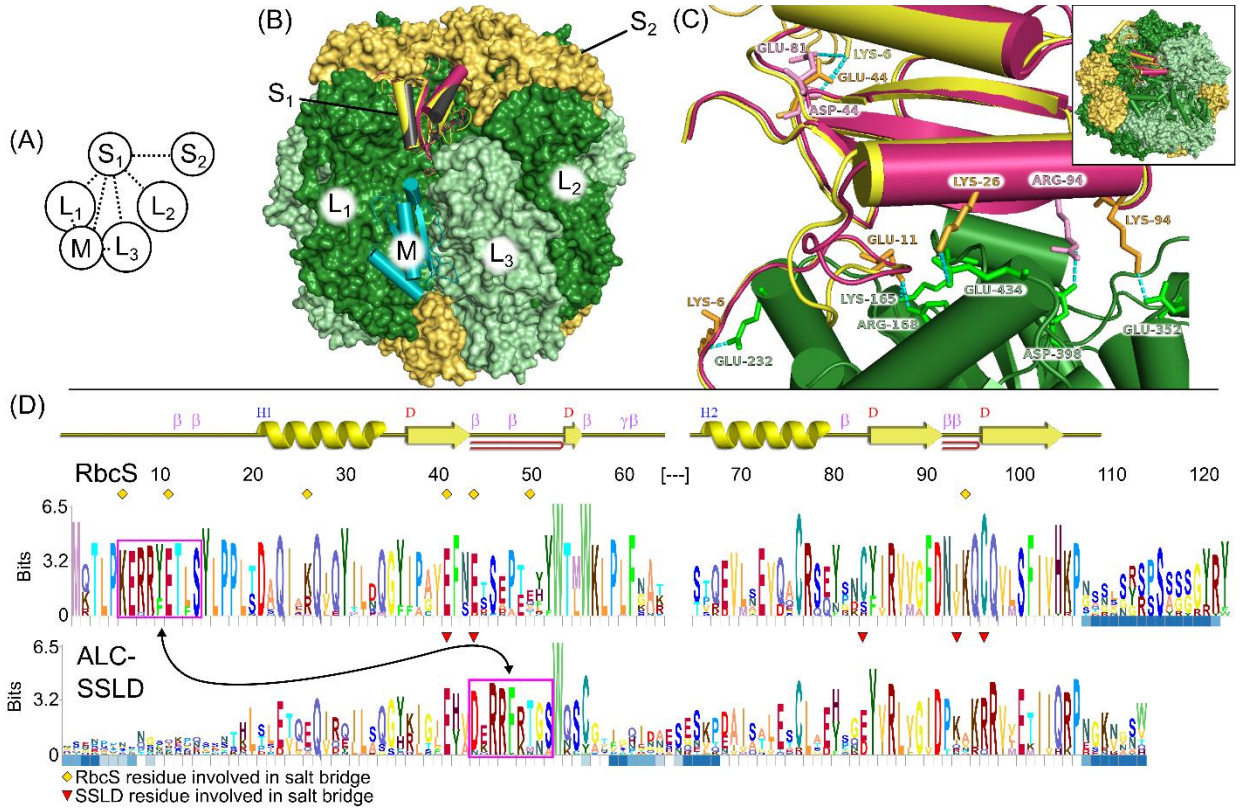


Figure 4.1: Structural comparison of RbcS and SSLD. (A) Schematic diagram of the interactions a single RbcS (S_1) or SSLD (M) has in a rubisco holoenzyme. Numbering based on van Lun et al.[117]. **(B)** Alignment of the SSLD found in ALCs (black is based on an SSLD model, burgundy on an RbcS model) to the RbcS (yellow) or SSLD (teal) binding sites of a L_8S_8 rubisco holoenzyme. **(C)** Interaction between $RbcL_1$ (green) and RbcS (yellow) or SSLD (burgundy) in the context of the RubisCO holoenzyme. Labeled residues highlight residues predicted to interact in a salt bridge interaction (Table 4.2), with label color based on the subunit it is from. Numbering for the SSLD is based on the trimmed region beginning at residue 317 in the full length *Fd*ALC. Dashed cyan lines show the predicted interacting atoms. **(D)** HMM-Logo highlighting the areas of sequence conservation between RbcS and the SSLD of ALC across cyanobacteria containing both. A schematic for the secondary structure of the homology model for RbcS from *F. diplosiphon* is presented above the two logos. Blue squares below each residue depict regions with gaps in significant portions of the MSA. Predicted salt bridges formed by the RbcS subunit are represented by yellow diamonds, while those formed by the SSLD are represented by red arrowheads. Magenta boxes, connected by an arrow, indicate a motif found in both MSA. Yellow diamonds indicate the RbcS residues involved in salt bridge interactions in the *F. diplosiphon* homology model, while red triangles indicate ALC-SSLD residues involved in salt bridge interactions in the S_1 position in the homology model (Table 4.2).

RbcS interactions with its nearest RbcL ($RbcL_1$, Figure 4.1A) are substantial, burying ~1600 $Å^2$ of surface area, with five predicted salt bridges, 13 hydrogen bonds, and a free energy of solvation of -5.2 $kcal \cdot mol^{-1}$ (Table 4.1). The remaining interfaces bury less area and are driven by fewer bond interactions, but favorably contribute to the overall L_8S_8 rubisco (Table 4.1, Column 5 – *Fd*RbcS).

When RbcS is replaced with the SSLD-modeled *Fd*ALC SSLD, most salt bridges are lost (Table 4.1, Column 3 – *Fd*ALC SSLD; Table 4.2), as are many hydrogen bonds. This is particularly true at the L_1-S_1 interface, where the absence of the crucial N-terminal loop (residues 3-17) of RbcS in the SSLD accounts for 3 of the 4 missing salt bridges at this interface, as well as the significant reduction of major buried surface area[95,171]. SSLDs have two features that may play a role in this interaction. First, SSLDs have a poorly conserved flexible linker at their N-terminus that could be involved in non-specific interactions. Additionally, a portion of the N-terminal loop in RbcS involved in L_1-S_1 interactions is positionally displaced in the primary structures of SSLDs. (Figure 1D, magenta box)[116]. Notably, the structural position of this region corresponds to a helix in SSLD structures but a loop in RbcS[171,213]. This 'displaced motif' region is conserved and resembles the important lost motif of the N-terminus of RbcS but was not noted in Ryan et al.[171] nor Wang et al.[213], possibly because without significant backbone rearrangement, this motif is unlikely to be positioned to bind in the same way and its conservation could be attributed to its role in binding at the SSLD equatorial interface. Overall, our modeling with the truncated *Syn*CcmM SSLD template is consistent with the experimental observations that the SSLD structure has a significant loss of favorable binding interactions at the RbcS interface.

Table 4.1: Predicted small subunit interactions of rubisco from *F. diplosiphon* and *Syn*7942. Results from analyses using Profunc and Pisa webservers of protein-protein interactions in CcmM complexed with rubisco or a homology model of rubisco from *F. diplosiphon*. Models contained homology models of *Fd*ALC SSLD or RbcS from *F. diplosiphon* or CcmM SSLD1 from *Syn*7942, as indicated by the row labeled "Interacting Subunit". These subunits were aligned to the position indicated in row 2 (labels based on Figure 4.1A) in rubisco L_8S_8 structures from *F. diplosiphon* (columns 3-5, 8), *Syn*6301 (column 6), or *Syn*7942 (column 7). Column 2 shows the rubisco subunit interface with the target SSLD/RbcS. Column 8, in bold, shows crystal structure data from Wang et al.[213].

	Interacting Subunit	*Fd* ALC SSLD	*Fd*ALC SSLD (RbcS model)	*Fd* RbcS	*Syn* 7942 CcmM SSLD1	***Syn* 7942 CcmM SSLD1**	*Fd* ALC SSLD
	Position	S_1	S_1	S_1	S_1	**M**	M
Number of Salt Bridges	L_1	1	1	5	0	**1**	2
	L_2	0	1	2	0	**0**	0
	L_3	0	1	0	0	**1**	2
	S_1	N/A	N/A	N/A	N/A	**1**	1
	S_2	0	2	1	1	**0**	0
Number of Hydrogen Bonds	L_1	2	9	13	1	**1**	4
	L_2	0	4	6	1	**0**	0
	L_3	1	2	1	0	**4**	3
	S_1	N/A	N/A	N/A	N/A	**1**	2
	S_2	4	2	3	3	**0**	0
Number of Non-bond Contacts	L_1	17	108	194	7	**17**	34
	L_2	2	26	77	7	**0**	0
	L_3	16	29	44	10	**39**	21
	S_1	N/A	N/A	N/A	N/A	**22**	21
	S_2	19	12	20	20	**0**	0
Pisa Interface ($Å^2$)	L_1	380.3	1326.5	1591.5	218.5	**236.4**	448.9
	L_2	91.0	495.1	614.9	260.0	**0.0**	0.0
	L_3	313.2	359.4	467.9	232.4	**511.3**	479.4
	S_1	N/A	N/A	N/A	N/A	**240.2**	239.3
	S_2	173.5	240.7	261.4	249.6	**0.0**	0.0
ΔG (kcal·mol^{-1})	L_1	0.9	1.3	-5.2	-0.6	**-1.4**	0.7
	L_2	1.0	2.6	-1.0	0.3	**N/A**	N/A
	L_3	-1.9	-1.5	-5.9	-1.0	**1.1**	3.1
	S_1	N/A	N/A	N/A	N/A	**0.9**	-0.7
	S_2	-3.1	-1.2	-3.2	-3.1	**N/A**	N/A

Table 4.2: Comparison of predicted salt bridges between RbcL and RbcS or the SSLD. Hypothetical salt bridges in the *F. diplosiphon* rubisco homology model interacting with either RbcS or *Fd*ALC SSLD models at position S_1 (see Figure 4.1A). The SSLD from *Fd*ALC modeled using either RbcS from *Syn*6301 as a template, or the CcmM SSLD1 from *Syn*7942 as the template. Residue numbering corresponds to *F. diplosiphon* sequences, with the SSLD corresponding to residues 317-424 of *Fd*ALC (e.g., residue R94 of the SSLD corresponds to residue R410 of *Fd*ALC).

Interaction	Rubisco:RbcS1	Rubisco:SSLD (RbcS Modeled)	RbcL:SSLD (ALC SSLD Modeled)
L_1-S_1	K165:E11	-	-
	R168:E11	-	-
	E232:K6	-	-
	E352:K94	D398:R94	D397:R94
	E434:K26	-	-
L_2-S_1	R188:E41	R188:E41	-
	K228:E50	-	-
L_3-S_1	-	D77:K91	-
S_2-S_1	K6:E44	K6:D44	-
	-	K6:E81	-

When the *Fd*ALC SSLD is modeled using RbcS as a template, part of the linker at the N-terminus of the SSLD (residues 1 – 17) is included in the model. Analysis of the *Fd*ALC SSLD in complex with rubisco suggested the potential for conservation of significantly more interactions (Table 4.1, Column 4 – *Fd*ALC SSLD [RbcS model]). Compared to the native RbcS, each interface buries slightly less area (~80% of that observed for RbcS) and is predicted to contain fewer hydrogen bonds and non-bonding contacts. Many salt bridges are potentially maintained or are similar for a total of 5 compared to the 8 found in RbcL-RbcS (Table 4.2). For example, while the SSLD model loses two salt bridges that contribute to the three structural checkpoints of the L_1-S interface described in van Lun et al.[117] (Figure 4.1C); this *Fd*ALC SSLD model is predicted to form a novel salt bridge with L_3 and K6 forms an additional salt bridge with S_2 instead of L_1. Although the SSLD of ALCs shows many regions of relatively low conservation, the regions that are important for RbcS interactions are generally well conserved even in the SSLD (Figure 4.1D), with the notable uncertainty of the displaced N-terminal motif

(magenta box) and linker. This suggests that when the flexible linker domain is also taken into account, *Fd*ALC SSLD could occupy an empty RbcS site.

4.5 Binding at the equatorial (M) position

As reported in Wang et al.[213], *Syn*CcmM SSLD1 forms favorable interactions with L_8S_8 rubisco in an equatorial position (we refer to this as position M). It forms a salt bridge with each rubisco subunit it contacts (L_1, L_3, & S_1) and forms some hydrogen bonds (Table 4.1, Column 7 – *Syn*7942 CcmM SSLD1). The ΔG at these three interfaces are less favorable overall than those calculated for the four interfaces of *Fd*ALC SSLD (Table 1, Column 3) and *Syn*CcmM SSLD1 in the S_1 position (Table 1, Column 6 – *Syn*7942 CcmM SSLD1). However, considering the presence of three salt bridges and that all RbcS positions were occupied in Ryan et al.[171], these data are consistent with the model that SSLDs would bind equatorially rather than displace an engaged RbcS subunit.

*Fd*ALC SSLD also shows potential for interaction with the equatorial position, although the ΔG values calculated from these preliminary models are relatively less favorable. When bound at M, it is predicted to form a greater number of salt bridges and more hydrogen bonds compared to the *Syn*CcmM SSLD1 (Table 1, Column 8 – *Fd*ALC SSLD).

4.6 Discussion

Here, we present a prediction that the SSLDs found in cyanobacteria may be able to substitute for RbcS in binding RbcL; we propose that this could occur in addition to recently demonstrated equatorial binding[171,213]. Our analysis considers the SSLDs found in the absolutely conserved

carboxysomal protein CcmM and the SSLDs found in the ALC, which is present in a subset of ecophysiologically diverse cyanobacteria. In the case of the *Fd*ALC SSLD, we found that its predicted binding with RbcL when substituting for the RbcS may be more favorable than that of the SSLD of CcmM of *Syn*7942, an organism that lacks the ALC. Although we find that the *Fd*ALC SSLD also could engage at the equatorial site, it may be a less favorable interaction than that observed for the SSLD of CcmM[171,213]. We suggest that the SSLD of *Fd*ALC has the potential for both equatorial binding and binding at the RbcS position. For *Syn*7942, the model organism used by Wang et al.[213] and more closely related to *Thermosynechococcus elongatus*[171], which also lacks an ALC, the predicted interface appears to point more favorably towards the equatorial binding found *in vitro*. Thus, it is possible that both the type of SSLD and the organism could influence whether the SSLD binds exclusively equatorially, especially for interactions with L_8S_8 rubisco holoenzymes. Additionally, in no case did we find that the SSLD bound better than the native RbcS, supporting the view that SSLDs cannot displace RbcS, though they might bind if sites are available.

In numerous quantifications of protein abundance in cyanobacteria, a shortfall of RbcS is found relative to RbcL[113,200]. This suggest that isoforms other than the L_8S_8 rubisco holoenzyme may be present *in vivo*. Indeed, experiments in *Syn*7942 report 5-6 RbcS for 8 RbcL, suggesting that RbcS binding sites could be available *in vivo*[113,200]. Additionally, we suggest it may be possible that the RbcL:RbcS ratio could be dynamically regulated with impacts on both enzyme activity and the binding position of RbcS. Moreover, rubisco undergoes numerous post-translational modifications that regulate its activity and subunit interactions (see review[67]). Phosphorylation can impact rubisco rate through phosphorylation sites at catalytic residues[111], and/or moderate

the interactions between RbcL, RbcS, & rubisco activase through partially unknown sites[1,70]. Phosphorylation reversibly alters the electrostatic interface, potentially affecting the favorability of equatorial driven SSLD interaction. These factors suggest mechanisms by which RbcS and SSLDs could be targets of dynamic regulation.

Throughout the evolution of diverse organisms containing Form I rubisco, a RbcL-RbcS fusion has never been observed. Given the importance of rubisco for survival, this is a significant clue to the potential for dynamic regulation of the RbcS:RbcL stoichiometry, potentially by the SSLDs. Though recent observations suggest that SSLDs bind primarily equatorially, we propose a potential for a dynamic relationship between multiple binding locations. Such dynamics could play a large role in the nucleation of carboxysomes, which is fascinating given the observed impact that SSLD-containing proteins have on carboxysome morphology[113,170]. These features would seem to depend heavily on the availability of RbcS binding locations, the flexibility of the protein structures, redox state, post-translational modifications, the species, the composition of the linker, the type of SSLD, and potentially even the subtype of RbcL[101], factors to consider in future investigations on the interaction between rubisco and SSLDs.

4.7 Acknowledgments

The authors would like to thank Dr. Sigal Lechno-Yossef for critical discussion of the data and interpretation of results. and Dr. Kaillathe “Pappan” Padmanabhan of the Macromolecular Computing Facility in the Michigan State University’s Department of Biochemistry and Molecular Biology for assistance with Rosetta installation and use of Linux clusters.

This work was supported by the U.S. Department of Energy (Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, grant no. DE-FG02-91ER20021 to B.L.M. and C.A.K.).

4.8 Author contribution

B.A.R. designed and conducted the research, analyzed and interpreted data, and wrote the article. C.A.K. and B.L.M. designed the research, analyzed and interpreted the data, and wrote the article.

CHAPTER 5

Linking the Dynamic Response of the Carbon Concentrating Mechanism to Carbon Assimilation Behavior in *Fremyella diplosiphon*

5.1 Abstract

Cyanobacteria utilize a CCM that enhances their carbon-fixation efficiency. Previous research has shown that the CCM is regulated by many environmental factors expected to impact photosynthesis. However, efforts to connect these findings to the functional effect on carbon assimilation rates are limited by the aqueous nature of cyanobacteria. Here, we present findings that utilize cyanobacteria in a semi-wet state on glass fiber filtration discs to establish carbon assimilation behavior using gas exchange analysis. In combination with qPCR and TEM analyses, we link the regulation of CCM components to corresponding carbon assimilation behavior in the freshwater, filamentous cyanobacterium *F. diplosiphon*. Inorganic carbon levels, light quantity, and light quality are all shown to influence carbon assimilation behavior. Results suggest a biphasic model of cyanobacterial carbon fixation. At low levels of CO_2, behavior is driven mainly by the cyanobacterium's C_i-uptake ability, while at higher CO_2 levels the carbon assimilation behavior is multi-faceted and depends on C_i availability, carboxysome morphology, linear electron flow, and cell shape. Carbon response curves utilizing gas exchange analysis offer rapid analysis of CO_2 assimilation behavior in cyanobacteria. These results provide an initial understanding important for modeling CO_2 assimilation in cyanobacteria and insight into how these data correlate to the stoichiometry of CCM components.

5.2 Introduction

The robust capability of cyanobacteria to fix carbon through photosynthesis is crucial to their key ecological role as one of Earth's major primary producers. Thus, understanding the regulation of their carbon fixation mechanisms is of general biological interest and a target of applied research on autotrophic production of biofuels and bioproducts. Cyanobacteria concentrate C_i through a

well-established CCM (see review[30]), which sequesters CO_2 and related enzymes and substrates into subcellular, proteinaceous, BMC called carboxysomes (see review[210]). As the carbon fixation steps of photosynthesis are often regulated to be kept in balance with the overall rate of photosynthesis[145], components of the CCM are likely to be tuned to environmental factors that affect photosynthesis as well. Indeed, both carbon transport and carboxysome components are upregulated in conditions where there is a greater need for C_i, such as growth under low CO_2 or HL[31,199]. Further study of the dynamic regulation of the CCM is poised to provide valuable insight into the modular components cyanobacteria utilize to coordinately control such features as the rate of photosynthesis and BMC morphology.

Uptake of C_i is the first component of the CCM. Since the cellular membrane is permeable to CO_2 but not HCO_3^-, cyanobacteria increase the flux of C_i into the cell by using HCO_3^- transporters as well as trapping CO_2 as HCO_3^- through the use of CO_2 hydrating enzymes. Constitutively expressed, active carbon transport[31,214] involves the low-affinity Na^+/ HCO_3^- symporter Bic in the cellular membrane[157] and the hydration of cytosolic CO_2 into HCO_3^- by NDH-1_4 (including subunits D4/F4/CupB) at the thylakoid membrane[184,224], thus driving HCO_3^- accumulation inside the cell. A parallel set of proteins with higher substrate affinity can be induced to increase C_i-uptake and includes: SbtA, another Na^+/ HCO_3^- symporter[183]; BCT1, an ATP-dependent HCO_3^- pump[144]; and NDH-1_3 (subunits D3/F3/CupA). Together, these complexes provide cyanobacteria with a high, and tunable, capacity for internal C_i influx as HCO_3^-, with some measurable leakage of CO_2 during C_i-uptake[208].

The second major aspect of the CCM is comprised of the protein shell of the carboxysome which forms a sub-cellular compartment that is permeable to HCO_3^- but not CO_2[46]. Both rubisco and CA are part of the carboxysomal cargo, which in conjunction with the high concentration of cellular HCO_3^-, drives the carboxylation reaction of rubisco forward with high local concentrations of its CO_2 substrate. In the case of β-carboxysomes, which are the type of carboxysomes formed in organisms with Type 1B rubisco, the *ccmKMNO* operon is crucial for carboxysome formation[156]. β-carboxysome biogenesis begins with rubisco aggregation by CcmM[38], a protein which contains an N-terminal, γ-class CA domain and 3-5 C-terminal repeats of an SSLD[112,114,116] that can interact with L_8S_8 rubisco and induce liquid-liquid phase separation[171,213]. Additionally, CcmM contains an internal ribosome entry site prior to the SSLD tail that results in expression of a truncated form that is crucial to carboxysome formation[113,114]. CcmN is then recruited to this condensate and, alongside full-length CcmM, interacts with at least CcmK2, the most abundant shell protein[38,44,94]. In most β-cyanobacteria, an additional β-class CA is recruited as well, though CcmM itself can fill this role in others[6,148].

The composition of the protein shell includes a diversity of paralogues with a potential range of substrate permeabilities. The faces of the shell are comprised of tessellating hexagonal proteins that contain a charged pore[92]. These hexagons result from repeats of a pfam00936 domain, either as hexamers of a single-domain monomer (BMC-H) or as trimers of monomers containing two domains in tandem (BMC-T). The BMC-H CcmK2 is present in all sequenced β-cyanobacteria and is usually joined by the highly similar CcmK1[192]. CcmK3 and CcmK4 always co-occur, have smaller pores, and seem to form heterohexamers and stacked dodecamers, suggesting mechanisms where they could dynamically alter the permeability of carboxysomes[193]. Other

paralogues include CcmK5 which appears to replace CcmK3/4, and CcmK6 which occurs in a heterocyst-forming clade of cyanobacteria. The BMC-T CcmO is essential for carboxysome formation[38,159], while the BMC-T CcmP forms stacked hexamers which may account for transport of larger substrates[36]. Separate from the paralogues forming the faces, CcmL forms pentamers that serve as the vertices of the intact carboxysome structure[203,206].

The myriad of components that comprise the CCM are responsive to environmental conditions. Both high light and low CO_2 levels tend to induce the expression of genes encoding many CCM components, especially for the high-affinity carbon transporters[31,76,124]. It has also been demonstrated that carboxysome morphology is dynamically responsive to light, C_i, and photosensory activity of cyanobacteriochromes[170,199,200]. However, many open questions remain in understanding how these changes control the carbon fixation capability of cyanobacteria.

There are many available methods for assaying carbon fixation and photosynthesis that have been applied to cyanobacteria. Measuring O_2 evolution, which probes linear electron flow at photosystem II (PSII) and shows reductions when CCM is compromised[116,154,198], via either mass spectroscopy or with an O_2 electrode is perhaps most common. Chlorophyll fluorescence can be used similarly but requires care in cyanobacteria to avoid phycobilisome absorbance or fluorescence confounding the results[39]. In more elaborate setups, carbon labeling can be used to determine rates of carbon assimilation and flux via either radiometry or mass spectroscopy. Due to the equilibration between CO_2 and HCO_3^-, both the media and cytosol can have stores of C_i separate from what is fixed, so the setup, lighting regiments, and measurement timing are used to distinguish between stores and assimilation of CO_2 and HCO_3^- [10,13]. In general, these techniques

are limited to end-point assays and/or are technically challenging. For terrestrial plants, a robust method derives net gas exchange from a plot of carbon assimilation vs intracellular CO_2 to establish steady state photosynthetic parameters non-destructively[55]. Carbon assimilation vs intracellular CO_2 curves are typically modeled with three distinct regions; low intercellular C_i values are limited by the reaction rate of rubisco, higher intercellular C_i values are limited by the rate of ribulose-1,5-bisphosphate regeneration (light limited), and at the highest intercellular C_i values the curves may show a saturation due to maximum utilization of triose phosphate pools[115]. Due to the aqueous nature of cyanobacteria and the slow, uncatalyzed equilibration with HCO_3^-, parallel methods have not been well established but are promising[141].

In this study, we analyze the carbon fixation characteristics of the filamentous, freshwater cyanobacterium *F. diplosiphon*. *F. diplosiphon* exhibits CCA, a process where cells respond to changes in the prevalence of light (primarily red vs. green in *F. diplosiphon* and many others) through altering the type and abundance of photosynthetic pigments, cell shape, and filament length[19,129]. Notably, cyanobacteriochrome (phytochrome-related) RcaE acts as a photoreceptor controlling CCA[23,89,189,209] and contributes to photoregulation of carboxysome morphology[170]. In order to probe the roles of CCA, RcaE, and carboxysome regulation in carbon assimilation (A, or the net rate of CO_2 uptake per unit area), we demonstrate that A can be measured using cyanobacteria in a semi-wet state with infra-red gas analysis of cyanobacterial discs. A is responsive to light quality and intensity, C_i availability, and the physiological state of cells. Generated carbon response curves (CRCs) establish that a Δ*rcaE* mutant strain exhibits impaired A under GL that is recovered under HL conditions. While A is only impaired in Δ*rcaE* relative to WT under GL, O_2 evolution is impaired in the Δ*rcaE* mutant under both GL and RL

conditions. Thus, we show that informative dynamic responses can be evaluated using CRCs in cyanobacteria and, together with measurements such as O_2 evolution, can be used to infer cellular propensity for C_i-uptake and active utilization in oxygenic photosynthesis. These findings are presented alongside TEM and CCM-related gene expression analyses in order to better understand how the regulation of CCM components contribute to the overall carbon fixation capabilities of cyanobacteria.

5.3 Results

5.3.1 Carbon assimilation measurements in *F. diplosiphon* respond to light, inorganic carbon availability, and physiological state

Glass fiber-filtered *F. diplosiphon* strains, i.e., *F. diplosiphon* discs, were analyzed using a LI-COR 6800 Photosynthesis System to detect CO_2 consumption. Carbon assimilation rates (A) in WT and Δ*rcaE F. diplosiphon* strains were responsive to light dosage (Q), showing light saturation at ~100 $\mu mol \cdot m^{-2} \cdot s^{-1}$, though this was higher (~150 $\mu mol \cdot m^{-2} \cdot s^{-1}$) in initial results with HL acclimated cultures (Figure 5.1A & Figure 5.2B). Thus, 300 $\mu mol \cdot m^{-2} \cdot s^{-1}$ was selected to serve as saturating light in further experiments. Under light saturating conditions, strains of *F. diplosiphon* exhibited a response in A to changing carbon levels in a standard CRC (Figure 5.1C-5.1F). Blank glass fiber filter discs wetted with fresh cell media were used as a control and show slightly negative A values that become more negative from 600-1000 ppm (Figure 5.2).

Typical A values ranged from 1-4 $\mu mol\ CO_2 \cdot m^{-2} \cdot s^{-1}$; this is in contrast to values of ~15-30 $\mu mol\ CO_2 \cdot m^{-2} \cdot s^{-1}$ seen in plants[55,115]. Samples were normalized by OD_{750} which had a roughly linear relationship with [Chl*a*] that varied somewhat with strain and growth condition (Figure

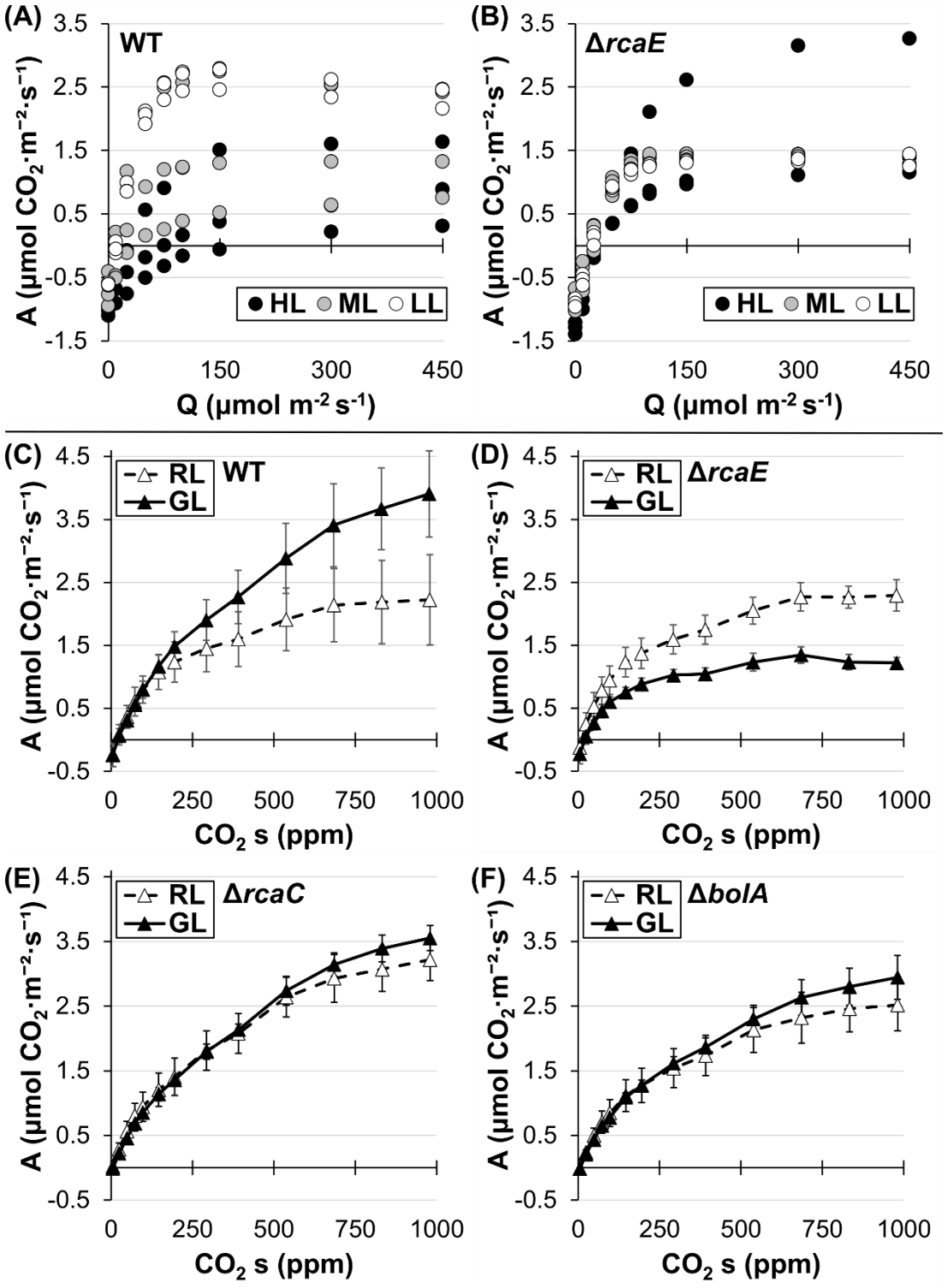


Figure 5.1: Carbon assimilation response to light and C_i availability. (A and **B)** Response to LI-COR chamber light at 400 ppm CO_2s for **(A)** WT and **(B)** Δ*rcaE F. diplosiphon* strains grown at low (12 µmol·m^{-2}·s^{-1}; white), medium (30 µmol·m^{-2}·s^{-1}; gray), and high (100 µmol·m^{-2}·s^{-1}; black) WL intensity in air, n = 3 for ML & LL and n = 1 for HL. **(C** to **F)** Response to supplied CO_2 at 300 µmol·m^{-2}·s^{-1} for **(C)** WT, **(D)** Δ*rcaE*, **(E)** Δ*rcaC*, and **(F)** Δ*bolA F. diplosiphon* strains grown at ~10-12 µmol·m^{-2}·s^{-1} red (white marks) or green (black marks) light conditions. Error bars represent 95% confidence intervals for n ≥ 5 from 2 independent biological replicates.

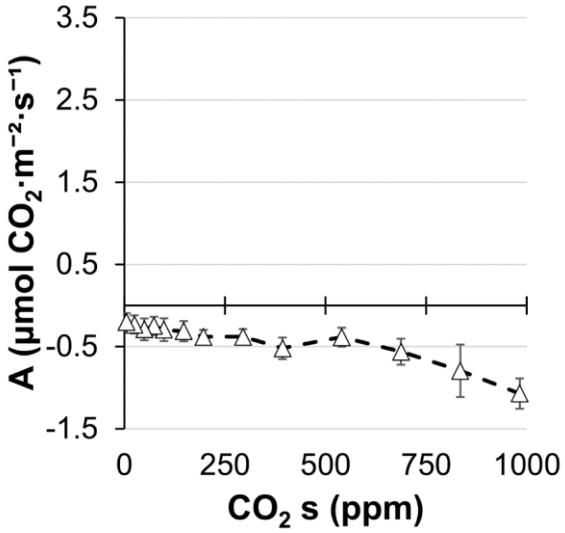


Figure 5.2: Carbon assimilation response to C_i availability for BG11/HEPES blank. Response to supplied CO_2 at 300 µmol·m^{-2}·s^{-1} for BG11/HEPES filtered through Whatman glass fiber filter paper as a blank for the semi-wet cyanobacteria gas exchange analysis. Error bars represent a 95% confidence interval from 3 replicates.

5.3). The volume of cells used resulted in ~10-55 µg Chl*a* that could be loaded in the gas exchange chamber. This is lower than chlorophyll levels used when assaying leaves, which range from ~120 – 180 µg in *Arabidopsis thaliana*[108,221] (incl. unpublished lab references) and likely contributes to the lower A values in *F. diplosiphon* than published plant values. Additionally, as the intercellular C_i flux in cyanobacteria is complex and not precisely modeled, response curves are presented with the [CO_2] levels in the sample chamber (CO_2s) as the independent variable. Lastly, compensation points (roughly the point where A becomes negative) measured in cyanobacterial CRCs appear to fall between 5-25 ppm CO_2s, which is likely lower than the typical values (25-100 ppm intercellular CO_2) found in higher order plants[28,115]. This observation is consistent with the presence of a CCM in cyanobacteria.

WT *F. diplosiphon* showed a difference in A only above 700 ppm CO_2 when comparing RL and GL acclimated cultures, with GL-grown cultures reaching higher A levels (Figure 5.1C). This result is consistent with previous findings measuring O_2 evolution, which show similar rates of

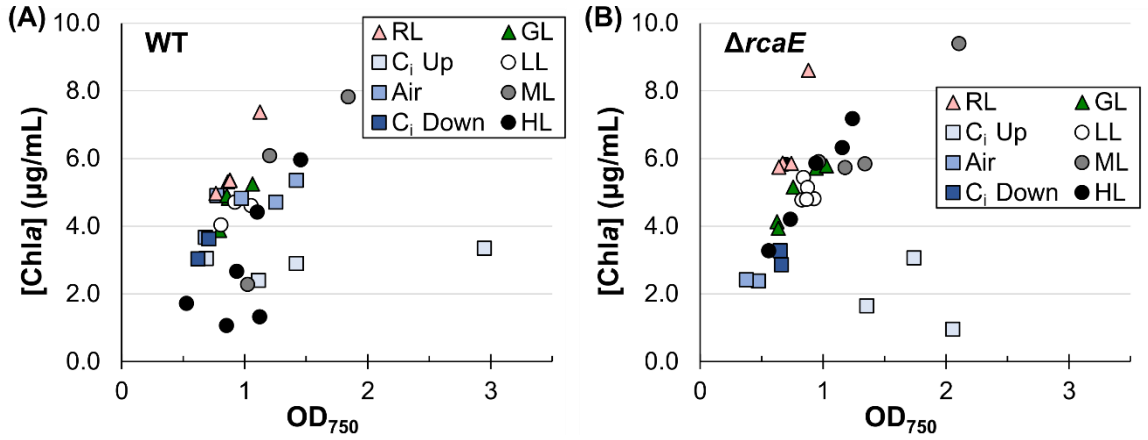


Figure 5.3: Chlorophyll *a* levels versus OD_{750} for cyanobacteria used in CRC analysis. Representative [Chl*a*] vs. OD_{750} measured in extracts harvested during CRC runs. Samples include **(A)** WT and **(B)** Δ*rcaE F. diplosiphon* strains grown under red (RL), green (GL), low (LL), medium (ML), or high (HL) white light conditions or in RL-enriched white light under air (Air), C_i upshift (C_i Up), or C_i downshift (C_i down), as described in the materials and methods section.

O_2 evolution between low intensity RL and GL at ambient CO_2[39]. By comparison, the Δ*rcaE* mutant demonstrated hampered A only under GL, with max A dropping from ~4.0 to ~1.3 $\mu mol \cdot m^{-2} \cdot s^{-1}$, whereas A in Δ*rcaE* was statistically indistinguishable from WT under RL (Figure 5.1C & 5.1D).

In order to test whether the difference in the net rate of CO_2 uptake is correlated with cellular pigmentation, we measured A in a Δ*rcaC* mutant strain, which lacks RcaC which is a DNA-binding regulatory protein that acts downstream of RcaE. This Δ*rcaC* strain has a constitutive GL-like CCA phenotype[71]. CRC analysis indicated no difference in A values for this strain between RL and GL, with values more similar to WT under GL (Figure 5.1E). This finding suggests that the GL physiological state is partially responsible for the higher A values.

In addition to pigmentation differences between RL and GL, *F. diplosiphon* exhibits cell shape differences controlled in part by RcaE[23] and its regulation of additional genes including morphogene *bolA*[189]. To separate potential impacts of the regulation of pigmentation from the

regulation of cell shape, we analyzed CRCs in a Δ*bolA* mutant strain, which has WT pigmentation but altered cell shape. A values in a Δ*bolA* mutant showed no difference between RL and GL and were closer to WT under RL (Figure 5.1F). Since the Δ*bolA* mutant loses the characteristic transition from a spherical cell in RL to a rod-shaped cell in GL, and shows A values more similar to RL, cell shape may also play a role in CRC behavior.

5.3.2 Effect of non-saturating light on carbon assimilation

In order to probe for light-limited regions of the CRC in cyanobacteria, we performed CRC analysis under non-saturating light conditions (25 and 50 $\mu mol \cdot m^{-2} \cdot s^{-1}$). We found that WT *F. diplosiphon* that had been grown in low light had near-saturating A values, even at as low as 50 $\mu mol \cdot m^{-2} \cdot s^{-1}$ (Figure 5.4A). However, A was severely impaired at 25 $\mu mol \cdot m^{-2} \cdot s^{-1}$ above 75 ppm CO_2s (Figure 5.4C). The Δ*rcaE* mutant also showed a decrease in A under non-saturating light conditions and was indistinguishable from WT at 25 $\mu mol \cdot m^{-2} \cdot s^{-1}$ (Figure 5.4B & 5.4D).

5.3.3 Effect of different light intensities during growth on carbon assimilation potential

Since high light is known to induce the components of CCM[31,76,199], we investigated the effect of growth under increasing light intensity on CRC behavior in WT and the Δ*rcaE* mutant. We utilized a multi-cultivator bioreactor system with green-enriched WL at low (LL; 12 $\mu mol \cdot m^{-2} \cdot s^{-1}$), medium (ML; 30 $\mu mol \cdot m^{-2} \cdot s^{-1}$) or high (HL; 100 $\mu mol \cdot m^{-2} \cdot s^{-1}$) light intensities. Growth rate increased as light intensity increased in both strains (Figure 5.5), although cells typically exhibited chlorosis ~7 d after induction of HL, indicating light stress. CRC analysis in WT indicated that responses to LL and ML were comparable, whereas HL showed a slight decline in A levels at high CO_2s (Figure 5.4E). In contrast, we observed increased A levels for

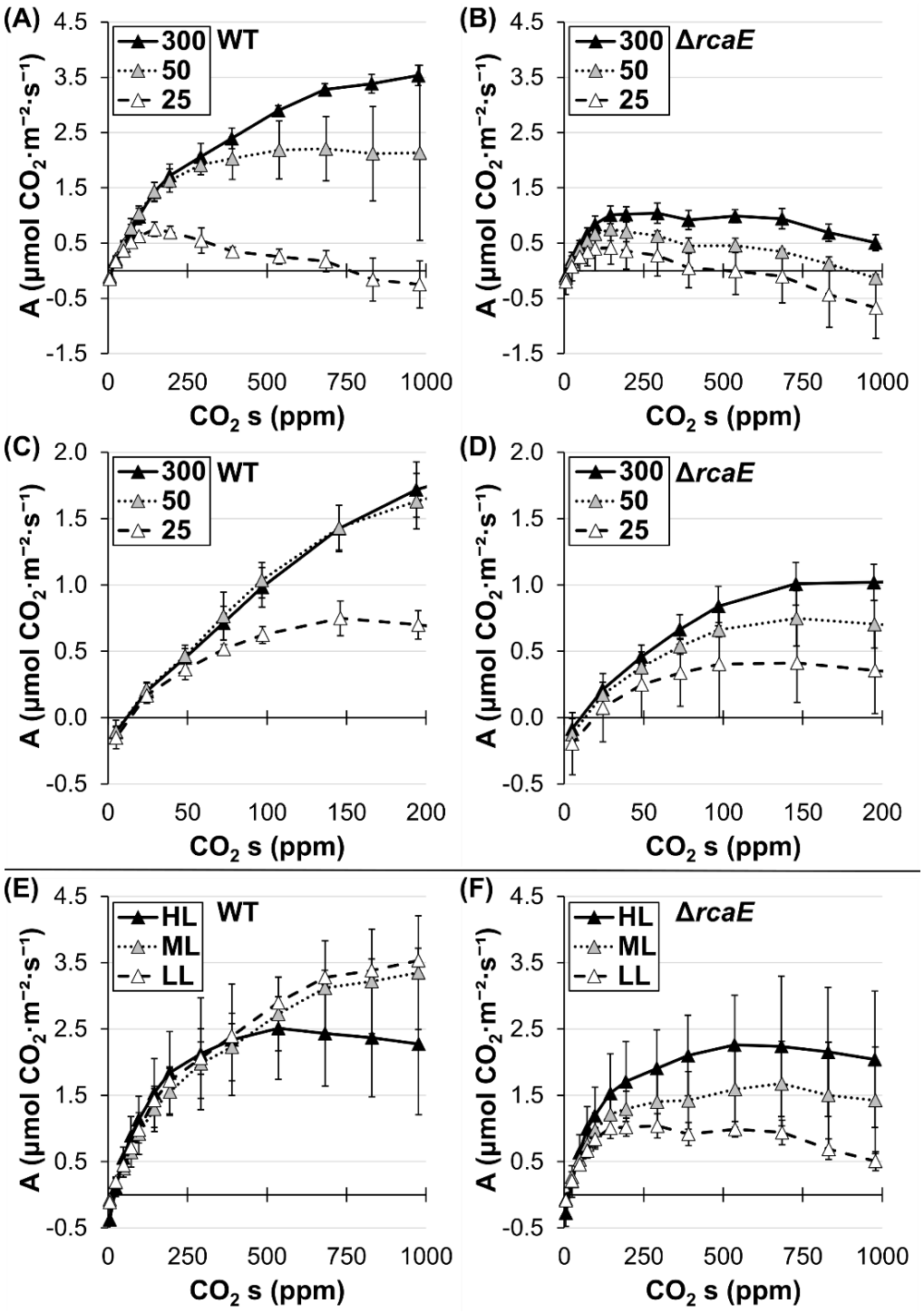


Figure 5.4: Carbon assimilation response to C_i availability in response to various light intensities. (A to **D)** Response to supplied CO_2 during runs at 300 µmol·m^{-2}·s^{-1} (black mark), 50 µmol·m^{-2}·s^{-1} (gray marks) or 25 µmol·m^{-2}·s^{-1} (white marks) for **(A** and **C)** WT and **(B** and **D)** Δ*rcaE F. diplosiphon* strains grown at low (12 µmol·m^{-2}·s^{-1}) GL-enriched WL. Panels C and D show 0-200 ppm CO_2s of panels A and B, respectively. Error bars represent 95% confidence intervals for n ≥ 3 from 2 independent biological replicates. **(E** to **F)** Response to supplied CO_2 at 300 µmol·m^{-2}·s^{-1} for **(E)** WT and **(F)** Δ*rcaE F. diplosiphon* strains grown at low (12 µmol·m^{-2}·s^{-1}; white mark), medium (30 µmol·m^{-2}·s^{-1}; gray mark), and high (100 µmol·m^{-2}·s^{-1}; black mark) GL-enriched WL intensities in air. Error bars represent 95% confidence intervals for n ≥ 4 from 2 independent biological replicates.

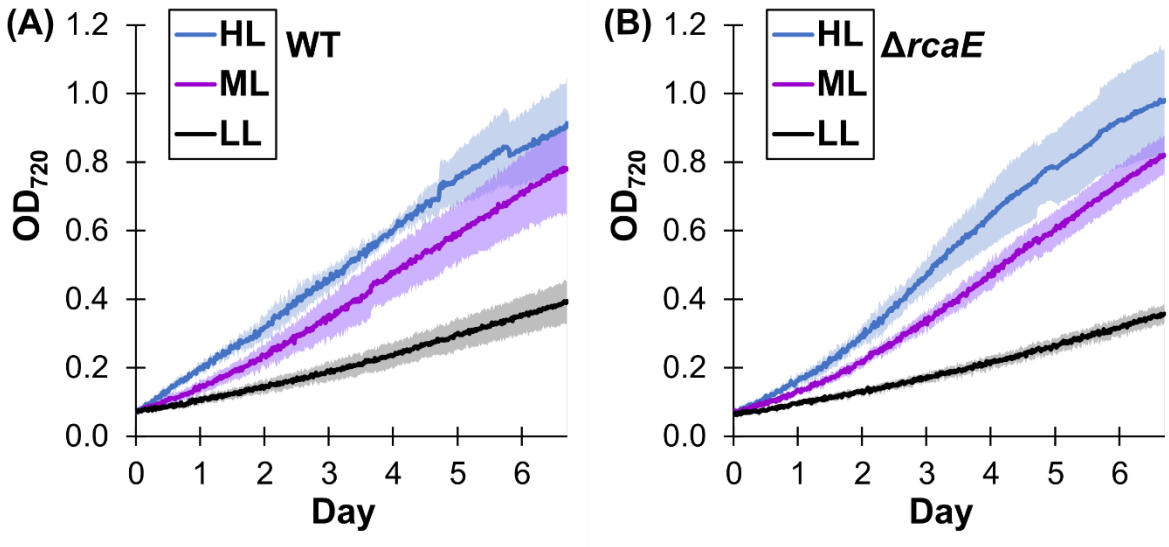


Figure 5.5: Growth rates of *F. diplosiphon* strains under increasing GL-enriched WL intensity. OD_{720} values versus time for **(A)** WT and **(B)** Δ*rcaE F. diplosiphon* strains grown under WL with dominant GL wavelengths at low (12 µmol·m^{-2}·s^{-1}; black line), medium (30 µmol·m^{-2}·s^{-1}; purple line) or high (100 µmol·m^{-2}·s^{-1}; blue line) intensity. Shaded area represents ± SD for n ≥ 4 from at least 2 independent biological replicates.

the Δ*rcaE* mutant in response to increased light intensity during growth, with the strain exhibiting near-WT levels under HL (Figure 5.4E & 5.4F). Also under HL acclimation, the two strains exhibited indistinguishable A values under non-saturating light conditions (Figure 5.6).Thus, it appears that WT utilizes light efficiently even under LL and that the Δ*rcaE* mutant is light starved under LL, with HL exposure resulting in recovery of its low A phenotype under green wavelengths.

5.3.4 Effect of inorganic carbon availability during growth on carbon assimilation

We next explored the impact of C_i availability on CRC behavior. Cells were grown in air, C_i upshift (3% CO_2), or C_i downshift (3 d growth in 3% CO_2 followed by a transfer to air for 19 h) in chambers illuminated with 35-40 µmol·m^{-2}·s^{-1} WL which induced growth consistent with growth in monochromatic RL and ambient air (Figure 5.1). Under these conditions, WT and

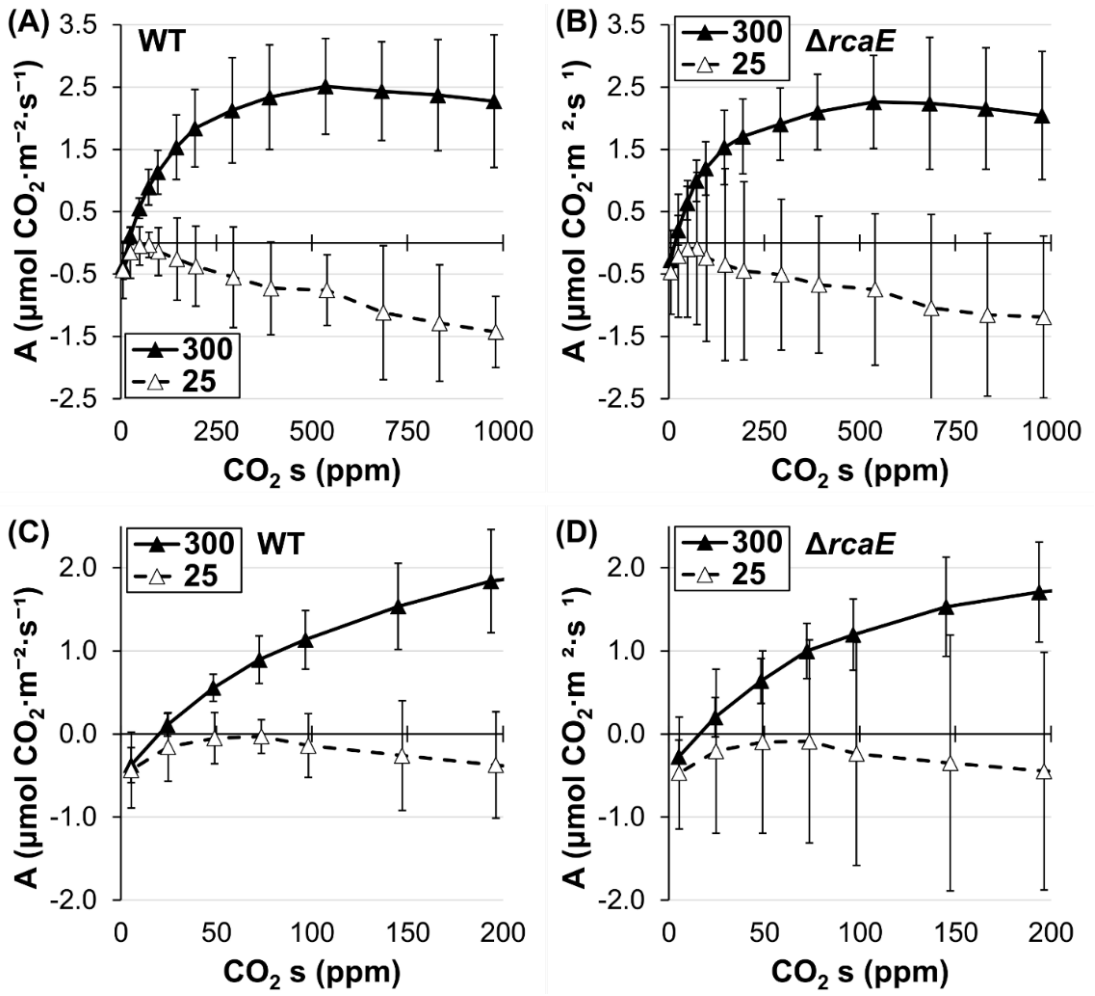


Figure 5.6: Carbon assimilation response to C_i availability under non-saturating light conditions after acclimation to HL. Response to supplied CO_2 at 300 µmol·m^{-2}·s^{-1} (black mark) or 25 µmol·m^{-2}·s^{-1} (white mark) for **(A** and **C)** WT and **(B** and **D)** Δ*rcaE F. diplosiphon* strains grown at high (100 µmol·m^{-2}·s^{-1}) WL intensity. Panels C and D show 0-200 ppm CO_2s of panels A and B, respectively. Error bars represent 95% confidence intervals for n ≥ 3 from 2 independent biological replicates.

Δ*rcaE* strains exhibited similar A behavior under air (Figure 5.7A & 5.7B). The behavior of these two strains was similar below 200 ppm CO_2s in all conditions, and as expected the compensation point appears to decrease as the cultures become more acclimated to lower C_i levels and induce high-affinity CCM systems (Figure 5.7C & 5.7D). During acclimation to C_i downshift, the two strains also performed similarly to each other in runs under non-saturating light conditions (Figure 5.8A & 5.8B); notably, non-saturating light shifted the apparent compensation point higher in both (Figure 5.8C & 5.8D). However, the Δ*rcaE* mutant strain

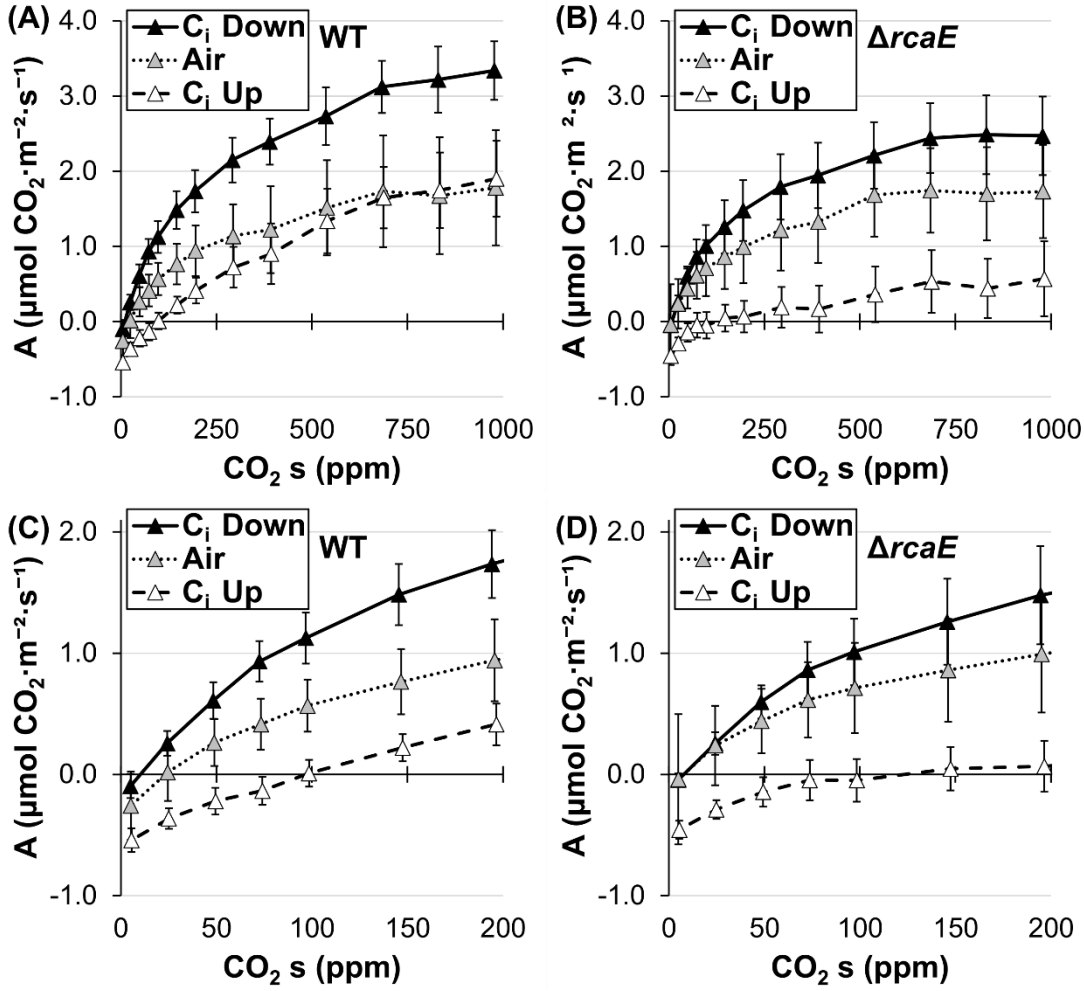


Figure 5.7: Carbon assimilation response to C_i availability after acclimation to various C_i levels. Response to supplied CO_2 at 300 μmol·m^{-2}·s^{-1}for **(A** and **C)** WT and **(B** and **D)** Δ*rcaE F. diplosiphon* strains grown at medium (~35 μmol·m^{-2}·s^{-1}) RL-enriched WL in air with 3% CO_2 enrichment (C_i Up; black mark), without enrichment (Air; gray mark), or under C_i downshift (C_i Down; white mark). Panels C and D show 0-200 ppm CO_2s of panels A and B, respectively. Error bars represent 95% confidence intervals for n ≥ 4 from 2 independent biological replicates.

exhibited a deficiency in A under C_i upshift and a less robust response to C_i downshift at higher CO_2s levels relative to WT (Figure 5.7B). Thus, the Δ*rcaE* mutant strain appears to have a deficiency in responding to C_i levels.

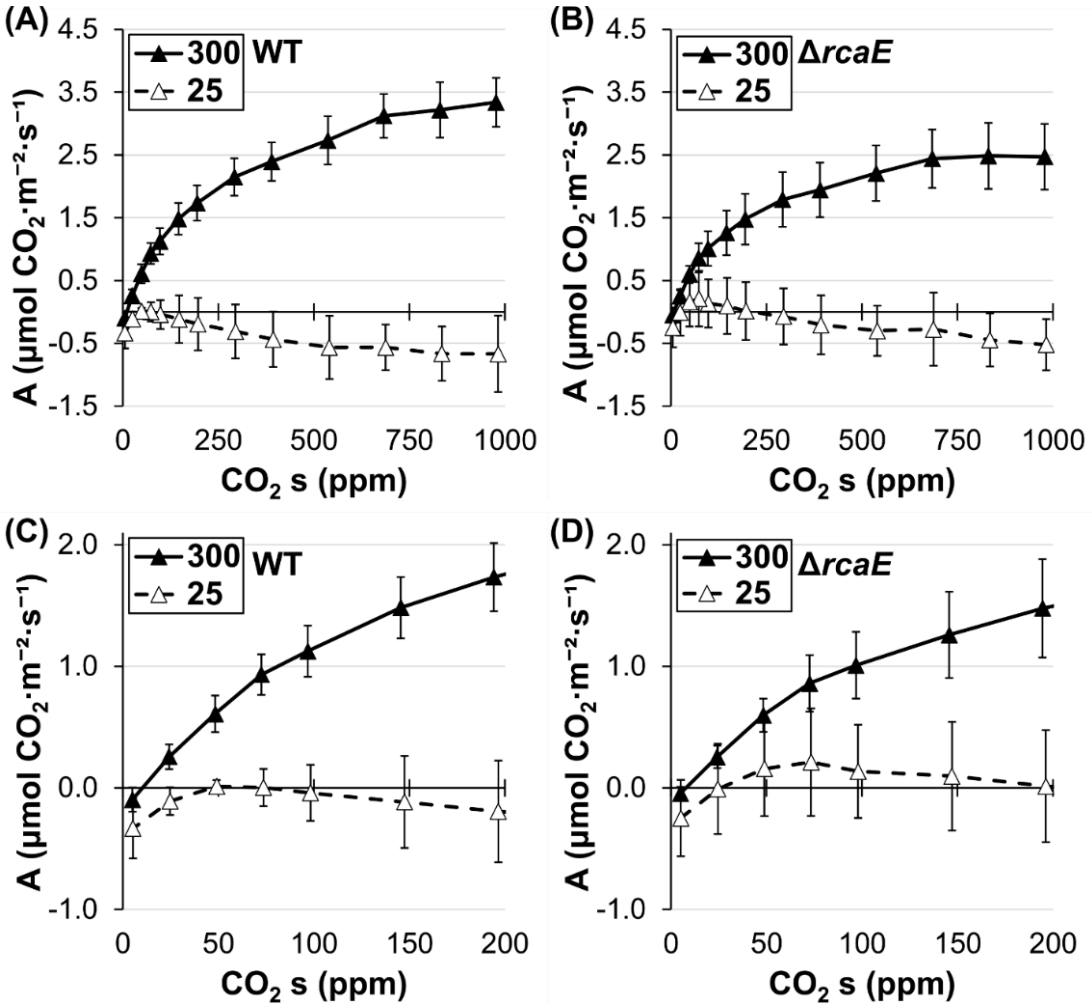


Figure 5.8: Carbon assimilation response to C_i availability in non-saturating light after acclimation to C_i downshift. Response to supplied CO_2 at 300 µmol·m^{-2}·s^{-1} (black marks) or 25 µmol·m^{-2}·s^{-1} (white marks) for **(A** and **C)** WT and **(B** and **D)** Δ*rcaE F. diplosiphon* strains grown at medium (~35 µmol·m^{-2}·s^{-1}) RL-enriched WL under C_i downshift. Panels C and D show 0-200 ppm CO_2s of panels A and B, respectively. Error bars represent 95% confidence intervals for n ≥ 3 from 2 independent biological replicates.

5.3.5 Rates of O_2 evolution in *F. diplosiphon* strains under red and green light

To compare our findings to established methods and to compare CO_2 uptake to active C_i utilization in oxygenic photosynthesis, we analyzed O_2 evolution in WT and Δ*rcaE* strains acclimated to RL or GL (Figure 5.9, white bars). We found that WT produced O_2 at higher initial rates in GL compared to cells grown under RL. O_2 evolution was significantly decreased in Δ*rcaE* under both RL and GL. Whereas CRC analysis only uncovered a defect under GL conditions, the Δ*rcaE* stain showed reduced O_2 evolution rates even when acclimated to RL.

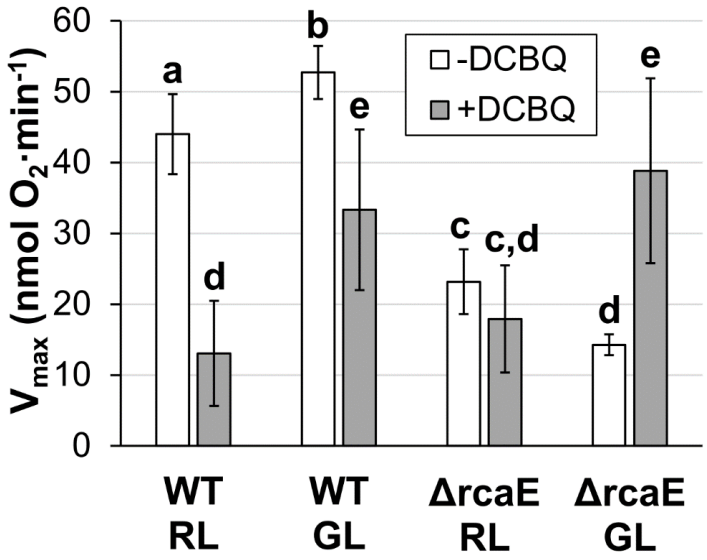


Figure 5.9: Oxygen evolution of *F. diplosiphon* strains acclimated to red or green light. O_2 levels measured after illumination by 250 μmol·m^{-2}·s^{-1} WL in *F. diplosiphon* strains (WT or Δ*rcaE*) grown under RL or GL, with or without the addition of 1 mM of the electron acceptor DCBQ. Error bars represent 95% confidence intervals for n = 5 (-DCBQ) or n = 3 (+DCBQ) from 2 independent biological replicates.

Additional experimentation utilized 1mM 2,6-dichloro-*p*-benzoquinone (DCBQ), which accepts electrons from PSII and acts as a photosynthetic uncoupler, in order to test for the total number of PSII centers capable of water oxidation[66,133]. In WT, we saw a substantial reduction in O_2 evolution proportional to the rates seen without DCBQ (Figure 5.9, grey bars). In contrast, addition of 0.5mM DCBQ in *Syn*6803 increased O_2 rates substantially[133]. The fact that the rates did not increase suggests that even without the addition of DCBQ, WT *F. diplosiphon* utilizes the majority of its PSII complexes that have sufficient excitement to split water (i.e., downstream regulation is not limiting WT). Furthermore, the decrease seen in A rates under RL (Figure 1C) may be attributable to PSII reaction rates as WT under RL exhibited lower O_2 evolution rates with and without DCBQ compared to cells in GL. Lastly, with the addition of DCBQ the Δ*rcaE* mutant showed no difference from WT in either light condition, due to a lack of response in RL and an increase in O_2 evolution in GL-acclimated Δ*rcaE* cultures (Figure 5.9). This finding

suggests that the apparent loss of photosynthetic rate in Δ*rcaE* under GL (as measured by both A and O_2 evolution rates) is not due to a deficiency in PSII complex rates.

5.3.6 TEM analysis of carboxysome morphology in response to light conditions and carbon availability

To contextualize the CRC behaviors, we analyzed carboxysome morphology in the studied conditions (Figure 5.10). We previously reported that under both RL and GL, carboxysomes were more abundant (higher number per cell section) and smaller in the Δ*rcaE* mutant relative to WT[170]. Additionally, carboxysome diameter decreased in both strains under GL and exhibited no light-quality-dependent changes in carboxysome abundance in either strain. We found that neither the Δ*rcaC* nor Δ*bolA* strains showed a difference in carboxysome size or shape between RL and GL (Figure 5.11A and 5.11B). These strains had significantly larger and fewer carboxysomes than in WT under GL, where WT exhibited a decrease in carboxysome diameter and trends towards higher carboxysome abundance.

Under increasing light intensity, WT showed a gradual increase in carboxysome diameter that was significant when comparing HL to LL ($p = 0.024$, Figure 5.11C), with no increase in carboxysome abundance (Figure 5.11D). The Δ*rcaE* mutant showed a similar increasing trend in carboxysome diameter with HL acclimating cultures showing a significant increase ($p < 0.001$ when comparing HL to either ML or LL, Figure 5.11C). In contrast to WT, the Δ*rcaE* mutant exhibited substantial increases in carboxysome number in response to increasing light. The Δ*rcaE* mutant did not exhibit its characteristic increase in carboxysome abundance until it was acclimated to ML or HL in these WL growth conditions relative to WT (Figure 5.11D).

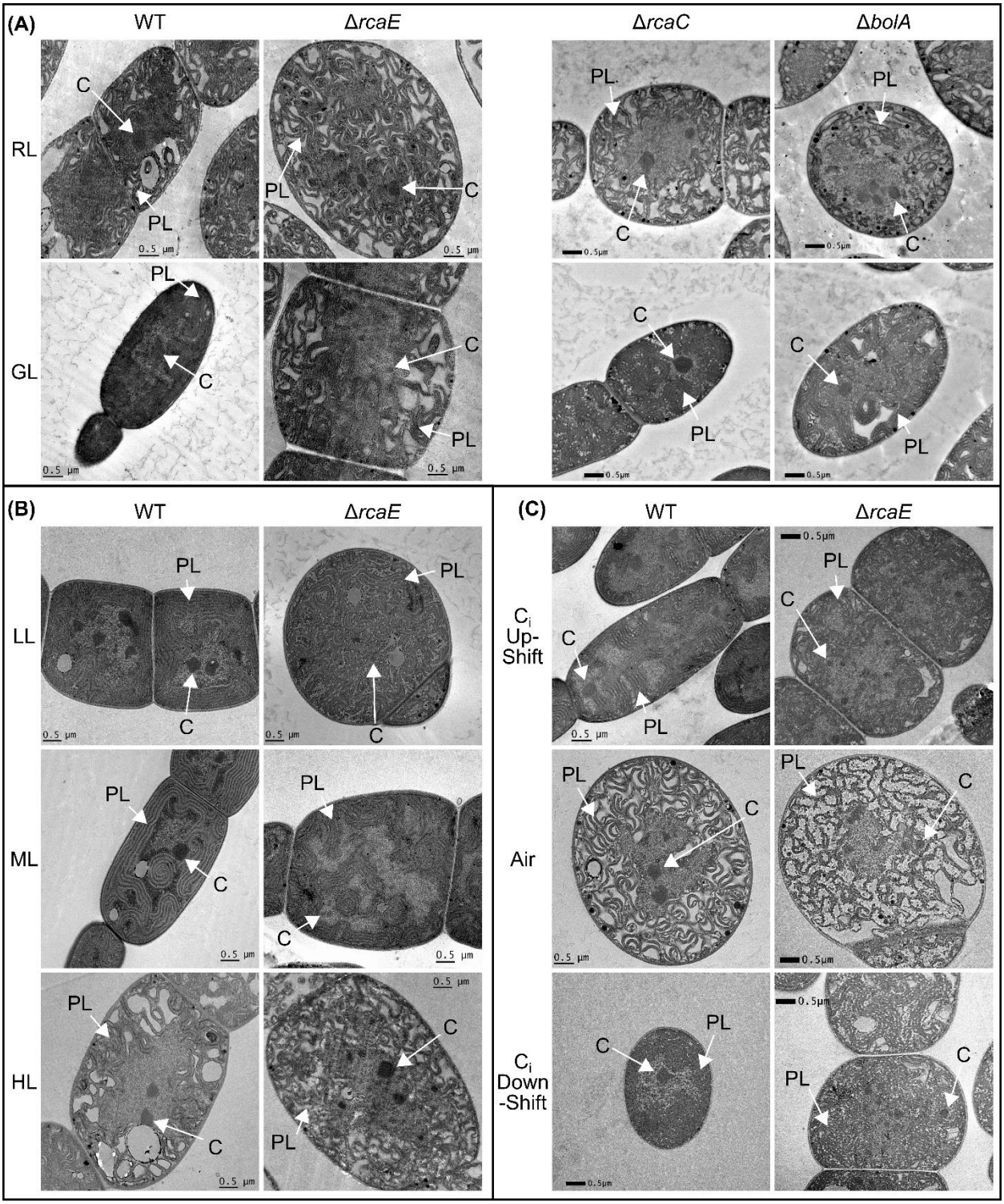


Figure 5.10: TEM analysis of cellular morphology of *F. diplosiphon* strains under changing light or C_i availability. Representative images of **(A)** WT, Δ*rcaE*, Δ*rcaC*, & Δ*bolA* strains under RL & GL, and WT & Δ*rcaE* strains under **(B)** increasing WL intensity or **(C)** decreasing CO_2 availability. Bars, 0.5 µm; C, carboxysomes; PL, photosynthetic lamellae.

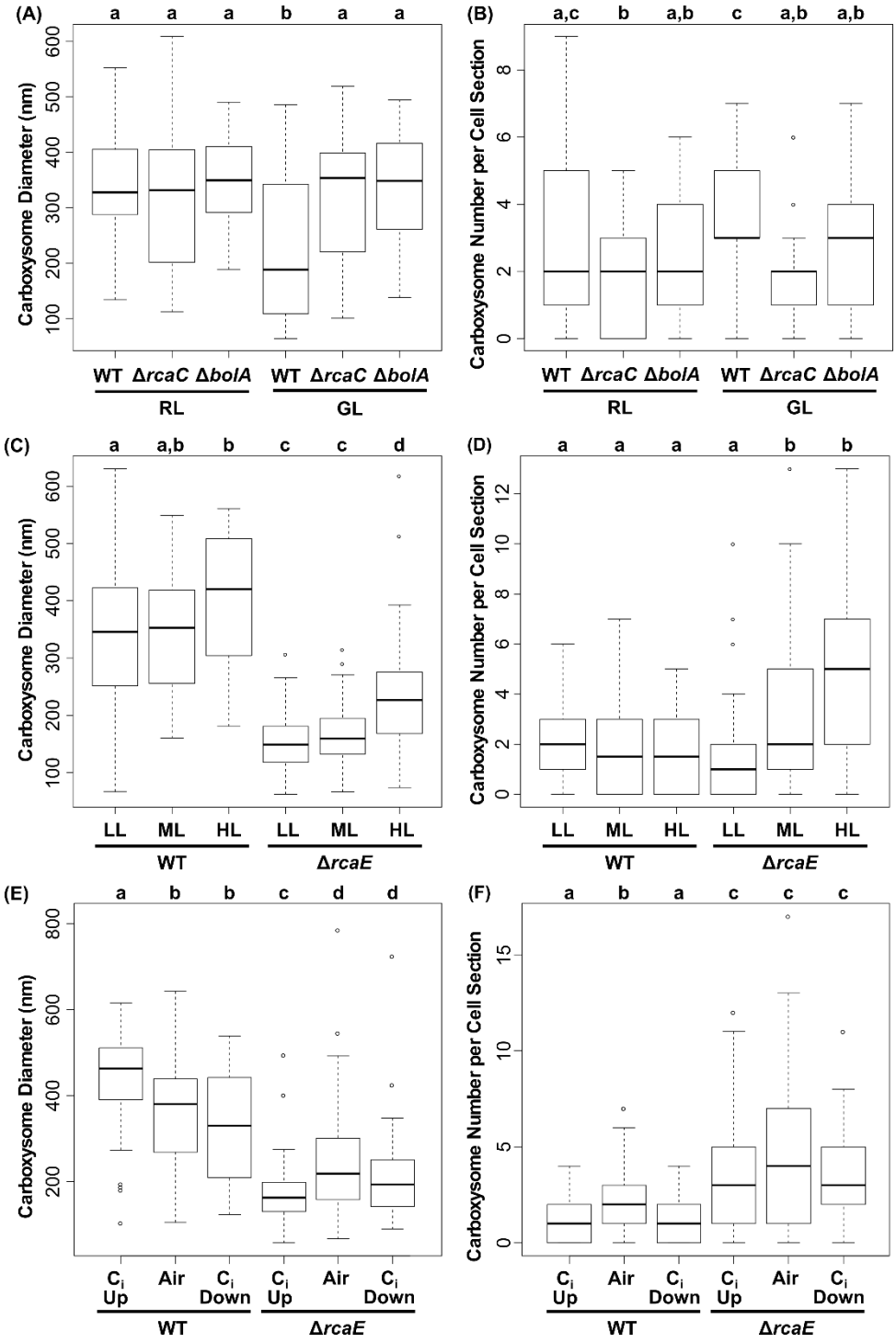


Figure 5.11: Carboxysome morphology under diverse physiological conditions. Boxplots displaying the full range of measurements of maximum carboxysome diameter and number of carboxysomes per cell section from TEM analysis for **(A** and **B)** WT, Δ*rcaC*, & Δ*bolA* strains of *F. diplosiphon* grown under RL and GL, and WT & Δ*rcaE* strains grown under increasing **(C** and **D)** WL intensity or **(E** and **F)** decreasing CO_2 availability. Lowercase letters indicate statistically significant groups ($p < 0.05$) within a panel. Corresponding average ± SE and sample size are presented in Table 5.1. Data for WT under RL and GL are reproduced here from Rohnke et al.[170] under the terms of the Creative Commons Attribution 4.0 International license, and data for WT under Air and C_i upshift are reproduced from Lechno-Yossef et al.[101] with permission.

Table 5.1: Quantification of average carboxysome sizes and numbers per cell section in Figure 5.11.

Condition[a]	Strain	Carboxysome size (nm)[b]	No. of carboxysomes/ cell section	Sample size (n) for carboxysome size measurements	Sections used (n) for no. of carboxysome/ cell section measurements
RL	WT	340 ± 19	3.0 ± 0.4	27	30
	Δ*rcaE*[c]	224 ± 12*	6.2 ± 0.6*	43	30
	Δ*rcaC*	323 ± 27	1.9 ± 0.3*	24	30
	Δ*bolA*	345 ± 15	2.5 ± 0.3	28	30
GL	WT	227 ± 19$^{\#}$	3.8 ± 0.3	45	30
	Δ*rcaE*	174 ± 5*,$^{\#}$	7.2 ± 0.9*	106	30
	Δ*rcaC*	325 ± 26*	2.0 ± 0.3*	18	30
	Δ*bolA*	336 ± 18*	2.6 ± 0.3*	31	30
LL	WT	318 ± 26	1.9 ± 0.3	26	30
	Δ*rcaE*	155 ± 9*	1.9 ± 0.4	34	30
ML	WT	354 ± 23	1.8 ± 0.3	21	30
	Δ*rcaE*	166 ± 7*	3.5 ± 0.6*,$^{\#}$	61	30
HL	WT	404 ± 25$^{\#}$	1.6 ± 0.3	19	30
	Δ*rcaE*	236 ± 11*,$^{\#}$	5.2 ± 0.6*,$^{\#}$	66	30
C_i Upshift	WT	436 ± 19	1.4 ± 0.1	42	60
	Δ*rcaE*	171 ± 7*	3.7 ± 0.4*	95	60
Air	WT	362 ± 15$^{\#}$	2.1 ± 0.2$^{\#}$	66	60
	Δ*rcaE*	244 ± 10*,$^{\#}$	4.5 ± 0.5*	134	60
C_i Downshift	WT	332 ± 27$^{\#}$	1.3 ± 0.2	22	30
	Δ*rcaE*	211 ± 14*,$^{\#}$	3.4 ± 0.5*	52	30

[a] Indicates conditions under which *F. diplosiphon* cells are grown as described in Methods and materials, i.e., RL, red light at ~10 to 12 µmol m^{-2} s^{-1}; GL, green light at ~10 to 12 µmol m^{-2} s^{-1}; LL, low GL-enriched WL at 12 µmol m^{-2} s^{-1}; ML, medium GL-enriched WL at 30 µmol m^{-2} s^{-1}; HL, high GL-enriched WL at 100 µmol m^{-2} s^{-1}; C_i Upshift, air enriched with 3% CO_2; Air, ambient air; C_i Downshift, growth under air enriched with 3% CO_2 followed by a shift to ambient air for ~19 h.

[b] Numbers for carboxysome size and carboxysome/cell section are represented as average ± SE.

[c] Statistical analyses, $p < 0.05$ indicated as follows: *, mutant vs. WT in same condition; $^{\#}$, significant difference vs. reference condition (RL, LL, or C_i Upshift) for same strain.

Ci availability also impacted carboxysome morphology. The WT strain exhibited the characteristic decrease in carboxysome abundance under C_i upshift (Figure 5.11F), while also showing an increase in carboxysome diameter (Figure 5.11E) (same data as reported in Lechno-Yossef et al.[101]). The C_i downshift conditions (transfer to air for ~19 h following growth in

elevated C_i) did not allow sufficient time for complete carboxysome acclimation, which is on the order of 2-4 d in *Syn*7942[199]. While the WT strain under C_i downshift showed similar carboxysome abundance to C_i upshift conditions, it had decreased size ($p = 0.003$), which could be part of the transition to the air-acclimated state (Figure 5.11E & 5.11F). The Δ*rcaE* mutant overall presented a misregulated response to C_i availability, showing a decrease in carboxysome diameter in response to C_i upshift (contrasted to the increase seen in WT, Figure 5.11E) and no significant response in carboxysome abundance (Figure 5.11F).

5.3.7 Transcriptional regulation of carbon concentrating mechanism components measured by qPCR analysis

Multiple components of the CCM are expected to be controlled at the transcriptional level in response to light and C_i availability[31,76,130,170,200]. For the studied strains and growth conditions, we analyzed the transcriptome with probes for CCM components (Table 5.2) using qPCR analysis. These analyses included probes for the carboxysome-related genes in the *ccmK1K2LMNO* and *ccmK3K4* operons, *ccmK6*, *ccmP*, *rbcL* or *rbcS* (rubisco subunits), *ccaA1/2* (carboxysomal CA), and *alc*. Genes related to C_i-uptake were also probed, including low-Ci induced *cmpA* (BCT complex), *sbtA*, and *ndhD4* (NDH-I_4 complex), constitutively expressed *ndhD3* (NDH-I_3 complex) and *bicA*, and a LysR-type transcriptional regulator with homology to *cmpR*[143] and *ccmR*[214], the latter two of which are both involved in the transcriptional response to C_i availability.

In an analysis of strains under RL and GL conditions (Table 5.3), Δ*rcaE* showed upregulation of *ccmM* and downregulation of *rbcS* under RL, whereas more significant changes were observed

Table 5.2: Primers used for qPCR probes in chapter 5.

Target Gene	Forward Primer 5'-3'	Reverse Primer 5'-3'
ccmK1	AACGAATTGGCAGGACATACT	GCAGGCGTAGAATCTGTGAA
ccmK2	AGGCTTGCACTTCCGATAC	TGCTGATGCGATGGTGAA
ccmK3	TGCTGCTGGAGAACAAGTAAA	GTAAAGTGGATCGGAAGGATGG
ccmK4	CAGGCAGTTGGAGCATTAGA	TCAGAAACATCGCCACGAATA
ccmK6	GAAGCAGTAGGACGAGTGAATG	ATTGGCGCTGCGATGAA
ccmL	GTCTACTCCTGCACCTACGATA	GTCTTCGAGGTGTGAAACTACTG
ccmM	GCAACAAGCTGACCGTTTAC	CTATCTGCAACGCACAAATATCC
ccmN	TGGCACTCAGATTTATGGTACAG	GTCCGAGATGGGTTCATTTAGAG
ccmO	CCATTACCTCCAAGCTCAGTAAA	CTCCTACCATCGCTGGAAATC
ccmP	TCATTCTAGCTCTCAAGGAGAAAC	CTAGAAACAACCCGAGGCTTTA
ccaA1	GCTCAAGTATACAGAGGCAACC	GAGTCAGTACATTCTCCGCAATAA
ccaA2	AACGAGCAGTTCGATTACCC	ATGCGCTCCCATTGTTCT
alc	CCGGCAACTATTCCTACCTTATC	TCGTGACAGGCAACGATTT
rbcL	GTTAGAAGGTGAGCGTGGTATC	GAAGCCCAGTCTTGGGTAAA
rbcS	TGTTCGGCGCTAAATCTACTC	GCTTGATGTTGTCAAAGCCTAC
LysR-Type	TCGGTCGGATTGCCTTTATTT	GCCGACAAGTAGCAAACAATTC
cmpA	CTGCATTAACCGCAGAGATTTG	GAGTATTGCTTTGGTGGCTTTG
sbtA	GTGGAACTGCGATCCGTAAT	ATGTATAGCGGGCGATGAATAC
ndhD3	TTCTCAGCGTTTCCCATCTC	CAGGTACGGTTGAGAAGAATCA
ndhD4	TGACTGCCGTGTACTTCTTAATC	GTAGGCGATCGCTCCAATATAC
bicA	GTTGCGGTTTGTACCGAATATG	TGTGGCTGTAAACCTGTGAG
orf10B	AGAACTACAGCGTCAGCTTAAT	CTGCTTCGCTTTCAGCATTT

under GL, particularly the downregulation of *ccmK3*, *rbcL*, *rbcS* and the low-C_i induced C_i-uptake genes relative to WT. The regulation of *ccmM*, *rbcL*, and *rbcS* were consistent with prior results[170], as were the downregulation of *sbtA* and *ndhD3* [130]. WT showed few differences between RL and GL conditions; *alc*, *bicA*, and *cmpA* were downregulated under GL. For many genes, Δ*rcaE* also exhibited downregulation under GL but with the magnitudes of change more extreme and more frequently statistically significant. The Δ*rcaC* mutant showed almost no difference between RL and GL, except a failure to downregulate *alc* under GL. Lastly, the Δ*bolA* mutant exhibited downregulation of *ccmK2*, *ccmK3*, *ccmK4*, and *sbtA* under RL.

Table 5.3: Relative expression of CCM genes in red versus green light conditions in *F. diplosiphon* strains. qPCR expression data of WT, Δ*rcaE*, Δ*rcaC*, and Δ*bolA F. diplosiphon* strains grown in RL or GL (~10-12 $\mu mol \cdot m^{-2} \cdot s^{-1}$) conditions.

	RL				GL			
	WT	Δ*rcaE*[b]	Δ*rcaC*	Δ*bolA*	WT	Δ*rcaE*	Δ*rcaC*	Δ*bolA*
n	5	5	6	6	5	5	6	6
ccmK1[a]	6.1 ± 0.4	6.7 ± 0.8	5.9 ± 0.5	5.7 ± 0.5	5.9 ± 0.4	6.0 ± 0.4	6.3 ± 0.5	6.2 ± 0.1
ccmK2	6.1 ± 0.4	6.8 ± 0.8	6.0 ± 0.4	5.5 ± 0.5*	5.8 ± 0.4	6.0 ± 0.2	6.2 ± 0.5	5.9 ± 0.2
ccmK3	5.1 ± 0.3	5.6 ± 0.5	5.0 ± 0.4	4.5 ± 0.3*	5.0 ± 0.2	4.4 ± 0.2*,#	5.1 ± 0.6	4.9 ± 0.2#
ccmK4	5.3 ± 0.3	5.5 ± 0.6	5.3 ± 0.4	4.7 ± 0.3*	5.0 ± 0.4	4.8 ± 0.4	5.6 ± 0.7	5.1 ± 0.3#
ccmK6	-0.5 ± 0.3	-0.3 ± 0.4	-0.4 ± 0.4	-0.1 ± 0.6	-0.7 ± 0.3	-1.0 ± 0.4#	-0.2 ± 0.6	-0.4 ± 0.1
ccmL	4.9 ± 0.6	5.5 ± 0.7	4.8 ± 0.4	4.3 ± 0.6	4.6 ± 0.3	4.4 ± 0.3#	4.9 ± 0.4	4.7 ± 0.2
ccmM	5.1 ± 0.5	6.3 ± 0.1*	5.2 ± 0.5	4.6 ± 0.6	5.2 ± 0.2	5.0 ± 0.3#	5.4 ± 0.5	5.0 ± 0.2
ccmN	3.8 ± 0.8	4.1 ± 0.5	3.4 ± 0.4	3.1 ± 0.6	3.5 ± 0.3	3.2 ± 0.2#	3.7 ± 0.3	3.5 ± 0.2
ccmO	3.8 ± 1.3	2.8 ± 0.7	3.5 ± 0.6	3.3 ± 0.8	3.5 ± 0.6	2.8 ± 0.2	3.6 ± 0.4	3.7 ± 0.1
ccmP	0.1 ± 0.5	0.5 ± 0.3	0.2 ± 0.2	0.2 ± 0.4	-0.2 ± 0.4	-0.4 ± 0.3#	0.2 ± 0.4	0.2 ± 0.1
ccaA1	-0.1 ± 0.7	-0.3 ± 0.7	0.5 ± 0.6	1.2 ± 1.2	-0.1 ± 0.2	-0.1 ± 0.4	0.0 ± 0.2	-0.1 ± 0.3#
ccaA2	-1.2 ± 0.7	-1.4 ± 0.6	-0.7 ± 0.3	0.1 ± 1.1*	-1.3 ± 0.4	-1.2 ± 0.5	-1.2 ± 0.3#	-1.0 ± 0.3
alc	3.0 ± 0.2	3.0 ± 0.6	2.7 ± 0.4	3.3 ± 0.5	2.2 ± 0.2#	1.7 ± 0.4#	2.6 ± 0.2*	2.9 ± 0.1*
rbcL	6.7 ± 0.9	6.0 ± 0.5	6.6 ± 0.7	5.5 ± 0.6*	6.4 ± 0.3	5.3 ± 0.3*,#	6.6 ± 0.5	6.0 ± 0.3
rbcS	7.1 ± 0.7	4.5 ± 0.1*	6.4 ± 0.6	5.8 ± 0.4*	6.8 ± 0.1	4.7 ± 0.2*	6.8 ± 0.3	6.2 ± 0.3*
LysR-Type	1.6 ± 0.5	1.3 ± 0.5	1.6 ± 0.4	1.4 ± 0.3	1.3 ± 0.4	1.5 ± 0.3	1.7 ± 0.3	1.6 ± 0.1
cmpA	1.8 ± 1.5	0.3 ± 0.5	0.0 ± 0.9	0.5 ± 0.6	-1.4 ± 0.4#	-2.3 ± 0.5*,#	-1.0 ± 0.6	-1.2 ± 0.3#
sbtA	3.9 ± 0.7	4.3 ± 0.5	4.1 ± 1.1	2.8 ± 0.6*	4.5 ± 0.7	2.5 ± 0.4*,#	4.3 ± 0.6	4.0 ± 0.5#
ndhD3	3.9 ± 0.5	3.9 ± 0.8	4.0 ± 0.6	3.5 ± 0.2	4.1 ± 0.2	2.7 ± 0.4*,#	4.3 ± 0.4	4.3 ± 0.4#
ndhD4	2.7 ± 0.3	3.2 ± 0.6	2.5 ± 0.4	2.2 ± 0.5	2.6 ± 0.2	2.6 ± 0.4	2.6 ± 0.3	2.6 ± 0.3
bicA	0.4 ± 0.3	0.5 ± 0.6	0.4 ± 0.4	0.8 ± 0.6	0.0 ± 0.2#	0.1 ± 0.3	0.3 ± 0.3	0.3 ± 0.3

[a] Data for each gene presented as $-\Delta C_q \pm$ SD relative to endogenous control gene *orf10B*, thus represent a $\log_2$ scale.
[b] Statistical analyses, $p < 0.05$ indicated as follows: *, mutant vs. WT in same condition; #, GL vs. RL in same strain.

Under increasing light intensity (Table 5.4), WT experienced significant upregulation of select HCO_3^- transporters (likely due to an increased linear electron flow), *ccmN*, and *ccmO*, alongside a downregulation of *rbcS* (possibly from degradation following high light stress). The Δ*rcaE*

Table 5.4: Relative expression of CCM genes under increasing light intensity. qPCR expression data of WT and Δ*rcaE F. diplosiphon* strains grown at LL (12 μmol·m^{-2}·s^{-1}), ML (30 μmol·m^{-2}·s^{-1}), and HL (100 μmol·m^{-2}·s^{-1}) GL-enriched WL intensity.

	LL		ML		HL	
	WT	Δ*rcaE*[b]	WT	Δ*rcaE*	WT	Δ*rcaE*
n	4	4	3	3	5	5
ccmK1[a]	6.7 ± 0.4	6.9 ± 0.8	7.1 ± 0.3	8.2 ± 0.3*,#	6.1 ± 0.8	7.1 ± 0.9
ccmK2	6.6 ± 0.6	6.9 ± 0.8	6.8 ± 0.2	7.7 ± 0.4*	5.9 ± 0.8	6.9 ± 0.9
ccmK3	5.6 ± 0.7	5.2 ± 0.7	5.7 ± 0.4	5.7 ± 0.4	4.7 ± 0.5	5.2 ± 0.5
ccmK4	5.5 ± 0.2	5.3 ± 0.5	5.8 ± 0.2	5.9 ± 0.4	5.3 ± 0.5	5.7 ± 0.7
ccmK6	-0.5 ± 0.2	-0.8 ± 0.4	-0.5 ± 0.2	-0.4 ± 0.2	-0.4 ± 0.4	0.1 ± 0.3#
ccmL	5.4 ± 0.5	5.3 ± 0.7	6.0 ± 0.5	6.6 ± 0.1#	4.8 ± 0.7	5.7 ± 0.8
ccmM	5.7 ± 0.4	5.7 ± 0.8	6.4 ± 1.1	7.2 ± 0.2#	5.1 ± 0.5	6.2 ± 0.5*
ccmN	3.7 ± 0.2	3.9 ± 0.7	3.9 ± 0.3	4.7 ± 0.4	6.9 ± 1.1#	5.3 ± 0.8*,#
ccmO	3.5 ± 0.2	3.3 ± 0.3	3.9 ± 0.1#	3.7 ± 0.2	7.8 ± 1.1#	6.4 ± 1.1#
ccmP	0.4 ± 0.2	0.3 ± 0.3	0.6 ± 0.1	0.7 ± 0.2#	0.1 ± 0.5	0.5 ± 0.4
ccaA1	0.3 ± 0.6	0.0 ± 0.7	0.0 ± 0.5	-0.3 ± 0.7	0.3 ± 1.2	0.1 ± 0.5
ccaA2	-1.0 ± 0.5	-1.4 ± 0.4	-1.0 ± 0.1	-1.2 ± 0.5	-0.4 ± 1.5	-0.6 ± 0.8
alc	2.6 ± 0.5	2.2 ± 0.7	3.5 ± 0.4	3.5 ± 0.3#	2.8 ± 0.8	3.3 ± 0.5#
rbcL	7.6 ± 0.7	6.4 ± 1.0	8.5 ± 1.7	8.0 ± 1.2	7.1 ± 1.1	7.0 ± 1.0
rbcS	7.9 ± 0.2	5.0 ± 0.3*	8.0 ± 0.9	5.4 ± 0.1*	6.6 ± 0.4#	4.5 ± 1.4*
LysR-Type	2.0 ± 0.4	1.8 ± 0.6	2.4 ± 0.4	2.5 ± 0.4	4.1 ± 0.2#	3.1 ± 0.4*,#
cmpA	-1.6 ± 0.5	-1.8 ± 0.8	-0.6 ± 0.3#	-0.4 ± 0.5#	5.8 ± 0.3#	4.3 ± 1.1*,#
sbtA	4.5 ± 0.2	3.5 ± 0.8	5.1 ± 0.6	4.3 ± 0.5	6.2 ± 0.2#	5.9 ± 0.7#
ndhD3	4.0 ± 0.6	3.3 ± 1.0	5.0 ± 0.5	4.6 ± 0.3	4.5 ± 0.3	5.0 ± 1.1
ndhD4	2.6 ± 0.6	2.9 ± 0.8	3.3 ± 0.3	3.7 ± 0.4	2.7 ± 0.2	3.2 ± 0.8
bicA	-0.6 ± 0.8	-0.2 ± 0.7	0.5 ± 0.2	0.3 ± 0.7	0.7 ± 0.4#	1.6 ± 0.6*,#

[a] Data for each gene presented as $-\Delta C_q$ ± SD relative to endogenous control gene *orf10B*, thus represent a $\log_2$ scale.
[b] Statistical analyses, $p < 0.05$ indicated as follows: *, Δ*rcaE* vs. WT in same condition; #, significant difference vs. LL in same strain.

mutant showed the characteristic downregulation of *rbcS* seen in other conditions. Additionally, it exhibited an upregulation of *ccmK1* and *ccmK2* under ML and *ccmK6* under HL that correlates with the increase in carboxysome abundance (Figure 5.10B, Figure 5.11D). Δ*rcaE* showed a similar upregulation of HCO_3^- transporters, *ccmN*, and *ccmO*, though not quite to the same extent

as WT. Lastly, Δ*rcaE* showed significant upregulation of *alc* in contrast to non-significant increases seen in WT. Since this protein is important for response to C_i limitation[101], this upregulation could be meaningful.

Both WT and Δ*rcaE* strains demonstrated significant differential expression of CCM components under decreasing C_i availability (Table 5.5). WT showed a general downregulation in shell proteins, *rbcL*, *rbcS*, and *ccmM* under C_i downshift, consistent with findings in *Syn*6803[48,214] and *Syn*7942[179]. It is interesting to consider how these data correlate with increased carboxysome abundance under C_i downshift reported in the literature[48,127,199] and this study (Figure 5.11E & 5.11F; C_i upshift vs. Air). As previously reported, *alc* is downregulated under C_i downshift and was reported to be involved in the decreased carboxysome abundance under C_i upshift[101]. Consistent with expectations, WT also exhibited significant upregulation of the low-C_i induced C_i-uptake genes. Additionally, as in HL, *ccmN* and *ccmO* were upregulated alongside the low-C_i genes rather than consistent with their operon. This high divergence of *ccmN* and *ccmO* is not seen in HL nor C_i upshift conditions in *Syn*7942[200] nor *Syn*6803[76,214]; in these strains the two genes do not diverge from the rest of the *ccm* operon. Moreover, *ccmO* is even located in a satellite operon in *Syn*6803 yet does not show this same pattern.

While WT upregulated the low-Ci induced C_i-uptake genes in both Air and C_i downshift, Δ*rcaE* only did so under C_i downshift. However, *ccm* gene transcription in the Δ*rcaE* strain was similar to WT under both C_i upshift and downshift overall, with the major differences being when the two strains transition from one to the other. Notably, Δ*rcaE* also recovered near-WT levels of *rbcS* under C_i downshift, possibly explaining the strain's recovery of A seen in this condition.

Table 5.5. Relative expression of CCM genes under decreasing carbon availability. qPCR expression data of WT and Δ*rcaE F. diplosiphon* strains grown under C_i Upshift (3% CO_2), Air, or C_i Downshift (19 h after a transfer from 3% CO_2 to air).

	C_i Upshift		Air		C_i Downshift	
	WT	Δ*rcaE*[b]	WT	WT	Δ*rcaE*[b]	WT
n	5	4	5	5	4	5
ccmK1[a]	6.9 ± 0.6	7.7 ± 0.4	6.1 ± 0.3[#]	6.7 ± 0.6[#]	6.2 ± 0.3	6.8 ± 0.5[#]
ccmK2	7.0 ± 0.5	7.6 ± 0.3	6.0 ± 0.3[#]	6.5 ± 0.5[#]	5.8 ± 0.3[#]	6.4 ± 0.5[#]
ccmK3	5.4 ± 0.3	5.2 ± 0.5	4.2 ± 0.3[#]	4.9 ± 0.1*	4.1 ± 0.2[#]	4.2 ± 0.5[#]
ccmK4	5.8 ± 0.3	5.6 ± 0.5	4.8 ± 0.2[#]	5.3 ± 0.2*	4.7 ± 0.2[#]	4.6 ± 0.4[#]
ccmK6	0.3 ± 0.3	-0.1 ± 0.3	-0.4 ± 0.4[#]	0.0 ± 0.3	-0.7 ± 0.5[#]	-0.4 ± 0.4
ccmL	5.7 ± 0.3	6.4 ± 0.4*	4.9 ± 0.2[#]	5.4 ± 0.5[#]	4.8 ± 0.2[#]	5.3 ± 0.4[#]
ccmM	6.5 ± 0.2	7.1 ± 0.5	5.1 ± 0.2[#]	5.9 ± 0.7[#]	4.7 ± 0.1[#]	5.3 ± 0.6[#]
ccmN	3.7 ± 0.3	4.8 ± 0.2*	5.4 ± 0.5[#]	3.8 ± 0.8*	7.4 ± 0.5[#]	7.2 ± 0.8[#]
ccmO	2.8 ± 0.4	3.9 ± 0.3*	6.2 ± 0.5[#]	2.7 ± 0.8*[,#]	8.0 ± 0.4[#]	7.8 ± 0.8[#]
ccmP	0.8 ± 0.3	1.0 ± 0.5	-0.1 ± 0.5[#]	0.6 ± 0.3*	-0.9 ± 0.3[#]	-0.7 ± 0.7[#]
ccaA1	-1.1 ± 0.2	-0.8 ± 0.2	-0.5 ± 0.4[#]	-0.3 ± 1.1	-0.9 ± 0.5	-0.4 ± 0.5
ccaA2	-1.9 ± 0.4	-1.8 ± 0.2	-1.4 ± 0.5	-1.3 ± 0.9	-2.2 ± 0.4	-1.7 ± 0.7
alc	3.8 ± 0.5	3.2 ± 0.6	2.5 ± 0.2[#]	3.3 ± 0.8	1.7 ± 0.3[#]	2.1 ± 0.6[#]
rbcL	8.3 ± 0.3	7.4 ± 0.3*	7.1 ± 0.6[#]	5.6 ± 0.8*[,#]	7.5 ± 0.4[#]	7.5 ± 0.5
rbcS	8.3 ± 0.5	5.6 ± 0.5*	7.5 ± 0.5[#]	4.2 ± 0.4*[,#]	7.6 ± 0.4	6.7 ± 0.3*[,#]
LysR-Type	2.1 ± 0.5	2.7 ± 0.2*	3.4 ± 0.2[#]	1.9 ± 0.7*	3.8 ± 0.3[#]	3.4 ± 0.2[#]
cmpA	-4.3 ± 0.4	-3.3 ± 0.4*	6.2 ± 0.2[#]	-1.9 ± 0.3*[,#]	6.6 ± 0.5[#]	6.7 ± 0.3[#]
sbtA	0.2 ± 0.3	0.2 ± 0.6	4.5 ± 0.3[#]	1.8 ± 0.2*[,#]	5.1 ± 0.6[#]	5.3 ± 0.3[#]
ndhD3	3.8 ± 0.3	3.8 ± 0.6	4.7 ± 0.3[#]	3.1 ± 1.1	5.4 ± 0.5[#]	6.1 ± 0.4*[,#]
ndhD4	3.0 ± 0.2	3.8 ± 0.3*	2.3 ± 0.2[#]	2.9 ± 0.8	2.6 ± 0.3[#]	2.9 ± 0.6
bicA	0.7 ± 0.2	0.7 ± 0.3	0.8 ± 0.3	1.4 ± 0.2*[,#]	0.5 ± 0.2	1.3 ± 0.3*[,#]

[a] Data for each gene presented as $-\Delta C_q \pm$ SD relative to endogenous control gene *orf10B*, thus represent a $\log_2$ scale.
[b] Statistical analyses, $p < 0.05$ indicated as follows: *, Δ*rcaE* vs. WT in same condition; [#], significant difference vs. C_i Upshift in same strain.

5.4 Discussion

5.4.1 Use of the carbon response curve in cyanobacteria

The CCM found in cyanobacteria has multiple modular components that can respond to dynamic environmental conditions and impact photosynthetic capacity. This capability is critical to

organismal fitness in the diverse habitats that photo-autotrophic cyanobacteria can grow. Many biological responses such as the upregulation of C_i-uptake genes under low C_i conditions are well established, and there has been an effort to cohesively model how the complex photosynthetic parameters of cyanobacteria arise from regulation of the CCM[121,141,200]. These efforts are generally limited to simple model cyanobacteria, and often serve poorly to rapidly measure net C_i consumption due to the aqueous nature of these organisms. Our work with *F. diplosiphon*, a freshwater filamentous cyanobacterium which undergoes CCA in response to light quality, highlights multilayered connections between CCM components, nutrient availability, and the physiological state of the cell[170]. Efficiently connecting these factors to overall carbon assimilation was a crucial step toward understanding how these organisms (and humans, as bio-prospectors) can optimize photosynthesis.

Utilization of gas exchange analysis to construct CRCs in cyanobacteria suggests that acclimation to dominant light quality through CCA has a nuanced impact on overall A behavior. WT *F. diplosiphon* cells assimilate more CO_2 when acclimated to GL despite having smaller carboxysomes and not being tuned to the red-enriched light of the LI-COR system (Figure 5.1C). The disruption of CCA through the loss of the major photoreceptor RcaE added additional layers of complexity; since RcaE influences the stoichiometry of carboxysome components and carboxysome size under both RL and GL[170], we expected a general decrease in net A. Instead, we found a GL-specific impairment (Figure 5.1D). While the small, more numerous carboxysomes of the Δ*rcaE* strain may contribute to overall A behavior, this observation cannot explain the higher A of WT under GL. These intriguing initial results prompted further exploration of the A behavior of cyanobacteria.

We provide evidence that physiologically relevant CRCs, similar in nature to the popular carbon assimilation vs intracellular CO_2 curves in plants, can be obtained from cyanobacteria in a semi-wet state using cyanobacterial discs. Cells show a dosage response to both light (Figure 5.1A) and CO_2, two major responses relevant to the development of advanced modeling of photosynthetic parameters in plants[115]. CRCs were also sensitive enough to show changes in apparent compensation point based on the physiological state of the cell (Figure 5.7C). Traditional O_2 evolution experiments show similar trends, with WT having higher rates under GL than RL while Δ*rcaE* exhibited higher rates under RL compared to GL (Figure 5.9). Despite this, the two methods differed when comparing WT and Δ*rcaE* strains under RL; Δ*rcaE* exhibited a decrease in O_2 evolution but not A rates, suggesting an impairment in utilization of CO_2 for oxygenic photosynthesis even though the two strains had similar CO_2 uptake rates under RL. Thus, CRCs of cyanobacterial discs offer novel insight into the CO_2 uptake behavior in cyanobacteria under a broad range of C_i levels. This method also significantly reduces the time required for equilibration between CO_2 and HCO_3^- which allows for dynamic responses to be studied. Thus, it is a promising technique that can both stand alone as a quick measurement of net carbon assimilation, and in conjunction with established systems that more deeply probe the HCO_3^-/CO_2 flux. In particular it serves to more directly test CO_2 utilization by cyanobacteria in contrast to well established procedures that test a cyanobacteria's utilization of HCO_3^-.

5.4.2 The low C_i phase of the carbon response curve (<100 ppm CO_2s) is C_i-uptake driven

Similar to the carbon assimilation vs intracellular CO_2 curve in higher-order plants, the CRC in cyanobacteria appears to have regions with different contributing factors. Though the different

regions do not yet appear to be as clear cut or as amiable to modeling as in plants, we analyze which factors appear to contribute to observed CRC behaviors. The presence of a distinct region at low ppm CO_2s is supported by the CRC behavior under non-saturating light (Figure 5.4A-5.4D). At very low light levels, A is mostly lost and instead the cyanobacteria appear to respire. However, there still seems to be a robust carbon assimilating behavior present upon reaching low C_i levels. This is likely due the low-Ci induced uptake system and reflects the high affinity cyanobacteria have for even trace amounts of C_i. Notably, this region is very robust and rarely exhibits differences; the Δ*rcaE* mutant is always indistinguishable from WT below 100 ppm CO_2s.

There are only two conditions in which we observed changes to this region. The slope and compensation point were very responsive to a culture's acclimation to different C_i availabilities with growth under C_i downshift prompting a robust A response even at minuscule C_i levels and a lowered apparent compensation point (Figure 5.7C & 5.7D). We were tempted to identify this as a 'light-independent' region, thus tested a hypothesis predicting that cultures acclimated to C_i downshift would not show a change in slope below ~100 ppm CO_2s even when analyzed under non-saturating light. This was not the case, however; non-saturating light reduced the A slope and increased the compensation point (Figure 5.8C & 5.8D). This observation suggests that light availability can affect this region, albeit only in specific conditions still related to C_i-uptake capacity. Thus, we propose identifying the low C_i region of the cyanobacterial CRC as a C_i-uptake driven region.

5.4.3 The high C_i phase of the carbon response curve (>100 ppm CO_2s) is responsive to multiple photosynthetic parameters

The identification of a C_i-uptake driven region at low C_i, separate from a region that reaches A_{max} at high C_i, suggests a biphasic CRC behavior reminiscent of those seen in C4 plants[85,150]. However, in contrast to C4 plants, which are mainly limited by phosphoenolpyruvate regeneration rate at high C_i, our results suggest that the high C_i region of cyanobacteria depends on many variables, including C_i availability, carboxysome morphology, linear electron flow, and cell shape.

Components of the CCM relating to C_i-uptake appear to have a broad effect on A behavior; upregulation of the low-C_i induced genes (Table 5.5) is correlated with an increase in A at all CO_2s levels (Figure 5.7A). Since this increase occurs in C_i downshift conditions where WT carboxysomes have not had sufficient time to acclimate to Air conditions (Figure 5.11E & 5.11F), this is one case where we can neatly attribute a change in A behavior directly to one major component of CCM. However, under HL, we see a similar induction of the low-Ci induced genes (Table 5.4) without the corresponding increase in A (Figure 5.4E). In the same vein, the increase in A under GL in WT (Figure 5.1C) is the opposite of what we would expect given the downregulation of *cmpA* seen under GL (Table 5.3).

Analysis of the Δ*rcaE* mutant strain provides some additional lines of inquiry that may offer insight. Unlike WT, increasing light intensity increased the mutant's A_{max} (Figure 5.4F). One possible explanation is that the mutant experiences a greater increase in overall carboxysome volume in response to HL (Figure 5.11C & 5.11D), perhaps evidencing a role of carboxysomes

in A behavior. It is also worth considering that the Δ*rcaE* mutant is likely less efficient at light utilization, as evidenced by the fact that addition of DCBQ as an electron sink increased the mutant's O_2 evolution rate in GL but not the WT's (Figure 5.9). Thus, the HL condition would be a greater boon to the mutant (as evidenced by its increase in A), but is likely stressful for the more efficient WT. This suggests that one or both of these factors (carboxysome size and linear electron flow) contribute to the determination of A_{max}. Secondly, the Δ*rcaE* mutant's behavior yields insights into the A phenotype of WT under GL. Though *cmpA* is downregulated under GL in WL, the Δ*rcaE* mutant shows much more significant downregulation of low-Ci induced genes (Table 5.3), which perhaps contributes towards the low A phenotype of Δ*rcaE* under GL. If this were the case, it would suggest that inducible C_i-uptake systems do indeed contribute, but are being masked in the high A phenotype of WT under GL.

Both Δ*rcaC* and Δ*bolA* mutants show few differences between RL and GL in the experiments performed in this study. Under both RL and GL, Δ*rcaC* that is constitutively in GL-like phenotypic state showed nearly identical A behaviors that were more similar to WT GL (Figure 5.1E), suggesting that GL acclimation may also contribute to WT's high A phenotype. Additionally, even though RcaC is downstream of RcaE, its loss did not seem to affect *ccm* gene expression while RcaE has major impacts (Table 5.3). As for the Δ*bolA* strain, it too showed nearly identical A behavior in both RL and GL but was instead more similar to WT under RL (Figure 5.1F). Since the Δ*bolA* mutant has WT-like pigmentation, this seems to rule out the pigmentation component of GL acclimation as a driving force for the high A phenotype. However, the Δ*bolA* mutant does have an enlarged, spherical cell shape under both RL and GL, so it is possible that the rod shape of WT *F. diplosiphon* cells under GL enhances C_i-uptake. As a

last point of consideration on the topic, WT has smaller, more numerous carboxysomes under GL than either the Δ*rcaC* or Δ*bolA* mutants, though it is difficult to conceive how the carboxysome data correlates with a high A phenotype in both WT GL and Δ*rcaC* under RL and GL.

Thus, we would conclude that A_{max} is a multifaceted region of the CRC in cyanobacteria that depends on C_i-uptake capabilities but is also impacted by at least one non-pigmentation factor related to GL acclimation such as higher PSII rates under GL (Figure 5.9), overall rate of linear electron flow, or cell shape. Additional factors may include temperature or metabolomics; it would be interesting to investigate both the flux through the Calvin-Benson cycle and triose-phosphate utilization in RL versus GL.

5.4.4 Regulation of the *ccm* operon in *F. diplosiphon*

Our findings that differential gene expression patterns of *ccmN* and *ccmO* are more similar to low-Ci induced genes than the rest of the *ccm* operon hold interesting implications. Though *ccmO* is only present in this operon in ~60% of cyanobacteria[192], this is still strong conservation, it is present in the *ccm* operon in *F. diplosiphon*, and *ccmN* is always conserved in this operon. Thus, it is noteworthy that we have identified an organism where these two genes are not co-transcribed with the rest of the operon. We have identified a LysR-type regulator immediately upstream of the *ccm* operon that is roughly co-transcribed with low-Ci inducible genes (Table 5.4 and Table 5.5). The predicted protein, fdiDRAFT81170, is the closest BLAST match for two notable low-C_i-induction related LysR-type regulators found in *Syn*6803: cmpR, a low-C_i induced regulator of *cmp*[143] and ccmR[214]. CmpR is particularly interesting, as it is found

upstream of *cmp* in *Syn*6803 and there is no LysR-type regulator near the *ccm* operon. As *F. diplosiphon* is missing a LysR-type regulator near *cmp*, it is possible that fdiDRAFT81170 fills a similar role. While we have not found a promoter region near *ccmN* or *ccmO*, it is possible that the co-regulation of these genes and the high affinity carbon transporters is due to the LysR-type regulator's novel location and associated function.

5.4.5 Impact

Through this study, we have integrated physiological analysis of the cyanobacterium *F. diplosiphon* with the novel application of gas exchange analysis to cyanobacteria. Like many cyanobacteria, *F. diplosiphon* performs CCA, which offers a useful system for studying the impact of light regulation, especially as it relates to photosynthesis. We have thoroughly explored the connection between the loss of RcaE, a cyanobacteriochrome that controls the CCA pathway, and the CCM. Analysis of CRCs provides a simple method to assay the carbon assimilation phenotype of cyanobacteria, connecting findings on how the stoichiometry of CCM components impact the structure and function of carboxysomes and C_i-uptake systems. Lastly, preliminary work to identify photosynthetic parameters identifiable through CRCs could contribute valuable insight into modeling and understanding the dynamic regulation of photosynthesis in cyanobacteria.

5.5 Materials and methods

5.5.1 Growth conditions

General culture inoculation and growth under RL and GL were performed as described in Rohnke et al.[170]. In brief, we utilized a short-filament strain of *F. diplosiphon* with wild-type

pigmentation identified as SF33 (WT)[42], an RcaE-deficient mutant strain (Δ*rcaE*) characterized by Kehoe and Grossman[89], an RcaC-deficient mutant strain (Δ*rcaC*) identified in our lab through a forward genetics screen, and a BolA-deficient mutant strain (Δ*bolA*) described in Singh and Montgomery[189]. Liquid cultures were inoculated from plated cultures and grown at 28°C under WL in BG-11/HEPES until they were diluted to an initial OD_{750} of 0.05 and transferred to experimental conditions.

The effect of light intensity was tested in a Multi-Cultivator MC 1000-OD system (Photon Systems Instruments, Drasov, Czech Republic) equipped with LED WL and autonomous monitoring of OD_{680} and OD_{720} according to the manufacturer's directions. Since the LED WL was GL dominant, starter cultures grown under GL were used for experiments involving the multi-cultivator to avoid WT showing a growth lag as it underwent CCA. Light conditions were set at a constant value of 12 $\mu mol \cdot m^{-2} \cdot s^{-1}$ (LL), 30 $\mu mol \cdot m^{-2} \cdot s^{-1}$ (ML), or 100 $\mu mol \cdot m^{-2} \cdot s^{-1}$ (HL). Since sustained HL conditions ultimately caused chlorosis, when high ODs were needed for harvesting for TEM and RNA extraction, ML and HL cultures were first grown at 12 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for 1-2 d prior to the onset of ML and HL conditions. Cultures grown this way were allowed to acclimate to the higher light intensity for at least 3 d prior to harvesting. All experiments involving HL grown cultures were harvested prior (within 6 d of HL onset) to the plateauing of OD that preceded substantial cell death.

The effect of carbon availability was tested in Multitron growth chambers (Infors HT, Bottmingen-Basel, Switzerland) at 30°C under WL (~35-40 $\mu mol \cdot m^{-2} \cdot s^{-1}$, RL-enriched) gassed with air either enriched with 3% CO_2 (C_i upshift) or unenriched (Air). To achieve C_i downshift,

we shifted cultures from C_i upshift to Air conditions after 3 d of growth and resuspension in BG11/HEPES lacking sodium bicarbonate, as described in Lechno-Yossef et al.[101] based on Wang et al.[214]. Cells were harvested for CRC, TEM, or qPCR analysis ~19 h after transfer to air (C_i downshift).

5.5.2 Carbon response curve analysis using *F. diplosiphon* discs

OD_{750} was measured in triplicate for cultures growing under the desired experimental conditions and were harvested between ODs of 0.6-1.2. A total volume equal to 11.8 absorbance units ($V = 11.8/OD_{750}$) was vacuum filtered through glass fiber filters with a pore size sufficiently small enough to capture >99% of *F. diplosiphon* cells (Whatman GF/A 47cm diameter, Sigma-Aldrich, St. Louis, MO), with a second layer of Whatman grade 1 filter paper to diffuse the filtrate more evenly. Disc diameter was selected to minimize extra surface area not necessary for the gas exchange chamber; about 47% of the disc surface area was exposed to the $6cm^2$ chamber and barely extended past the gaskets. Cyanobacterial discs were handled carefully with forceps, briefly dabbed on filter paper to remove excess wetness, kept on BG11/HEPES agar plates, and swiftly analyzed to minimize environmental perturbance. CO_2 levels were measured with infra-red gas analysis in a LI-COR Photosynthesis System 6800 (LI-COR, Lincoln, NE), with one end of a strip of damp Whatman grade 1 filter paper placed underneath the disc as a wick. The other end was submerged in ddH_2O to maintain disc dampness for the duration of the experiment, which was found to greatly increase the duration that steady state could be maintained to ~45 min (data not shown).

The chamber was illuminated by the standard "Sun+Sky" (RL-dominant) regime with a leaf temperature of 28°C, flow rate of 500 μmol s^{-1}, and source air with 12 ppm H_2O. For the standard CRC, the initial CO_2 supplied was 1000 ppm and the sample was allowed to equilibrate for at least 5m and until a steady state had been maintained for at least 3m.The CRC followed a gradient of 1000, 850, 700, 550, 400, 300, 200, 150, 100, 75, 50, 25, and 5 ppm followed by a return to 400 ppm with automatic infra-red gas analysis matching. The sample was allowed to equilibrate for ~2-3 min at each time point for a total run time of ~25 min after initial equilibration.

5.5.3 O_2 evolution analysis

O_2 evolution was measured using an Oxytrace+ O_2 electrode (Hansatech Instruments Ldt, Norfolk, England) illuminated by an acrylic projector bulb. Illumination was maintained at ~250 μmol·m^{-2}·s^{-1} measured with a LI-250 Light meter (LI-COR) equipped with a quantum sensor (model US-SQS/L, Heinz Walz CmbH, Effeltrich, Germany). Cells containing ~10 μg Chl*a* (based on OD_{750} extinction coefficients [Figure 5.3]) were harvested, washed 2x in 3mL BG11/HEPES lacking sodium bicarbonate, and ultimately resuspended in 1mL BG11/HEPES lacking sodium bicarbonate. Cyanobacteria were placed in the chamber and spiked with 1M sodium bicarbonate (Sigma-Aldrich) to a final concentration of 2mM prior to illumination. If applicable, 2,6-dichloro-*p*-benzoquinone (DCBQ; Sigma-Aldrich) was then added to a final concentration of 1mM. Cells were allowed to equilibrate at ambient light for ~1.5m then illuminated. The O_2 evolution V_{max} was recorded as the peak rate reached within 10m of illumination.

5.5.4 TEM analysis

TEM analysis was performed according to the methods of Rohnke et al.[170] for all experimental conditions. For C_i upshift and air conditions, 60 cell sections were randomly selected and analyzed for carboxysome number in WT and Δ*rcaE* with carboxysome diameters measured in 20 of these sections. In all other strains and conditions, 30 cell sections were analyzed with 10 analyzed for carboxysome diameter as well. Samples were prepared from at least two independent biological replicates. As a modification to the original method, some samples were analyzed using a JEM 1400 Flash TEM (JEOL USA Inc., Peabody, MA), still at an operating voltage of 100V.

5.5.5 qPCR analysis

The abundance of *ccmK1*, *ccmK2*, *ccmK3*, *ccmK4*, *ccmK6*, *ccmL*, *ccmM*, *ccmN*, *ccmO*, *ccmP*, *ccaA1*, *ccaA2*, *alc*, *rbcL*, *rbcS*, fdiDRAFT81170 (a LysR-type transcriptional regulator), *cmpA*, *sbtA*, *ndhD3*, *ndhD4*, and *bicA* transcripts were measured in relation to the internal control *orf10B* in total RNA extracts from *F. diplosiphon* strains under various experimental conditions as described previously[101,170] in accordance with MIQE guidelines [32]. In brief, this involved harvesting ~20 mL of exponentially growing cells upon reaching the target OD_{750} (~0.5-0.6), handling the samples on ice and flash freezing the cell pellet within 1 h of harvesting, extracting with Trizol reagent incubated at 95°C followed by wash steps, DNase treatment (TURBO DNA-free kit, Invitrogen, Madison, WI), and RNA quantification using a NanoDrop ND-1000 Spectrophotometer. Reverse transcription was performed using qScript cDNA SuperMix (Quantabio, Beverly, MA) kit and qPCR was prepared using Fast SYBR-Green master mix (Applied Biosystems, Foster City, CA) in 384-well plates (Applied biosystems) with a 10 μL

reaction volume, both according to the manufacturer's directions. Probe sequences are provided in Table 5.2. RNA quality was assayed using gel electrophoresis and genomic contamination was controlled for by verifying that no-template-control samples had C_q values greater than 5 cycles higher than the respective unknowns. Data reflect three technical replicates for each of at least three independent biological and are presented using the delta C_q method (ΔC_q) in order to foster analysis between several strains and conditions.

5.5.6 Chlorophyll extraction

Chl*a* was measured spectrophotometrically according to the methods of de Marsac and Houmard[122] as described for analysis in *F. diplosiphon*[187]. Samples were harvested in parallel with CRC analysis as a secondary validation of normalization by OD_{750} and at least three independent biological replicates were analyzed.

5.5.7 Statistical Analysis

Experiments were performed with $n \geq 3$ from at least two biological replicates for all experiments.

5.6 Acknowledgements

We are grateful to Dr. David T. Hanson of the Department of Biology at The University of New Mexico for innovating the utilization of filtered liquid cultures in photosynthetic gas exchange analysis and for his kind guidance during our troubleshooting of the method. We are also grateful to Dr. Thomas D. Sharkey and Dr. Berkley Walker of the Michigan State University (MSU) Plant Biology Laboratory for providing LI-COR 6800 availability and detailed discussion of

methodology and results. This study is indebted to the hard work of Kiara Rodriguez of the University of Puerto Rico, who we were fortunate enough to host for a productive Research Experience for Undergraduates opportunity. Her contributions appear in Figure 5.1A-5.1D and Figure 5.4. We are also thankful to Alicia Withrow of the MSU Center for Advanced Microscopy for her extensive assistance with the transmission electron microscope and providing a diamond knife to use for this study.

This work was supported by the U.S. Department of Energy (Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, grant no. DE-FG02-91ER20021 to B.L.M.). We also thank Melissa Whitaker (supported by National Science Foundation grant no. MCB-1243983 to B.L.M.) for strain maintenance and culture production.

CHAPTER 6

Conclusions and Perspectives

The work described in this dissertation has centered on the characterization of the regulation of the cyanobacterial CCM in response to dynamic environmental conditions expected to influence photosynthetic rate. I successfully characterized the tuning of the CCM and related carbon fixation potential to C_i and light availability in *F. diplosiphon*, which are trends also found in other cyanobacteria. I also have made significant contributions to understanding the regulatory roles of CCA and the rubisco activase homologue, ALC, in the CCM. These results support the view that the CCM of cyanobacteria can be finely tuned as a part of the regulation of photosynthesis. This regulation impacts the carbon assimilation behavior of *F. diplosiphon*, highlighting the functional impacts of regulating C_i-uptake and carboxysome morphology.

RcaE, a cyanobacteriochrome that controls the RL vs. GL response of CCA in *F. diplosiphon*, has an important impact on CCM components and assimilation behavior. Its loss led to a complicated change in the photosynthetic behavior of *F. diplosiphon*, including changes in pigmentation and reductions to O_2 evolution rates (Figure 5.9), demonstrating an impaired photosynthetic efficiency of a Δ*rcaE* mutant. This loss also resulted in a perturbed stoichiometry of CCM components related to the impaired photosynthesis and resulted in a distinct misregulation of carboxysome abundance and morphology. Carboxysomes were smaller and more numerous in a Δ*rcaE* mutant strain (Figure 2.2). Interestingly, this effect was not correlated with the function of known components of CCA regulation downstream of RcaE, with Δ*rcaC* and Δ*rcaF* mutants showing no changes to carboxysome morphology under RL, limited changes under GL (Figure 2.6, Figure 5.10A, Figure 5.11E & 5.11F), and only a few changes in *ccm* gene regulation in RL and GL (Figure 2.7, Table 5.3). These findings suggested that RcaE works through distinct effectors, or through effects on the physiological state of the cell, to regulate

carboxysome morphology by altering the ratio of carboxysome shell components to cargo components (Figure 2.3, Figure 2.4, Table 5.3). These changes were correlated with a decrease in carbon assimilation potential under GL (Figure 5.1C & 5.1D), showing a functional link between the detection of light quality, tuning of the CCM, and carbon assimilation potential.

Our increased understanding of the role of RcaE supports the rising view that regulation of the ratio of shell components to cargo components is a major driving factor of carboxysome dynamics. Larger ratios of shell:core generally lead to smaller and/or more numerous carboxysomes, and vice versa[113,200], and this includes an important consideration of the two forms of CcmM, with M35 (containing only SSLDs) being part of the core components while M58 recruits carboxysome shell components[114] (Figure 2.4). This understanding of the importance of carboxysome component stoichiometry on structure and abundance provides a general understanding that the expression levels of cargo intended for incorporation into a BMC must be controlled and offers a mechanism with which to regulate BMC morphology.

This research also provided evidence that ALC, the rubisco activase homologue found in cyanobacteria, should be considered as an important part of the carboxysomal cargo. In addition to being localized to carboxysomes (Figure 3.6), expression of the *alc* gene was found to correlate with carboxysome morphology across multiple conditions. *alc* was upregulated in response to C_i upshift (Figure 3.10), conditions under which carboxysomes are larger and less numerous (Figure 3.11), which is consistent with *alc* participating in the determination of the carboxysomal shell to cargo ratio as part of the cargo. An Δ*alc* mutant strain of *F. diplosiphon* lost the response to C_i upshift (Figure 3.11), implying a causal effect. In support of this view, *alc*

expression was correlated with carboxysome abundance under GL; WT under GL had smaller, more numerous carboxysomes than either Δ*rcaC* or Δ*bolA* mutant strains (Figure 5.11B, Table 5.1), with an upregulation of *alc* in both mutants relative to WT (Table 5.3). Compared to RL, WT under GL had smaller carboxysomes (Figure 5.11A, Table 5.1) and the downregulation of *alc* in GL was the only differential expression of the carboxysome-related genes apparent. Thus, while the function of ALC remains undetermined (Figure 3.7F), it appears to have an interesting role as carboxysomal cargo interacting with rubisco, potentially mediating rubisco network formation (Figure 3.7D, Figure 3.9) alongside CcmM-35[171,213], with potential impacts on rubisco packing density. In addition to expression level, ALC's interaction with rubisco appeared to be ATP-dependent in *F. diplosiphon* (Figure 3.7E), suggesting that ALC could contribute to regulation of carboxysome morphology in response to multiple factors.

These findings support further investigations into the regulation of carboxysome morphology in response to the cellular energy status. The misregulation of PPB in the Δ*rcaE* mutant (Figure 2.10) in parallel with the misregulation of carboxysomes implicates a potential functional connection between phosphate availability and carboxysomes. PPB abundance is inversely correlated with ATP levels, with PPB diameter increasing during dark periods in *Syn*7942 under light/dark cycles[181] when ATP levels are low[25,78]. Assessing ATP levels in WT *F. diplosiphon* under RL vs GL and contrasting these with levels in the Δ*rcaE* null mutant would test whether the smaller, more numerous PPB in Δ*rcaE* and WT under GL reflect changes in the ATP status of cells, which could draw further correlations to CCM regulation. If so, the role of ATP in determining carboxysome morphology could then be tested by experimentally reducing ATP levels through the use of energy transfer inhibitors[25].

Regulatory mechanisms controlling carboxysome morphology must also be considered alongside overall cell morphology. Recently, it was shown that carboxysome positioning can be driven by the intercellular oscillations patterns of McdA and McdB concentrations, which are also dependent on cell shape[118]. In light of these findings, it seems possible that localization of both PPB and carboxysomes are interrelated, potentially by this positioning system. Moreover, the impact of CCA's regulation of cell shape is thus a crucial point of consideration. Misregulation of PL also appears to impact carboxysome positioning, seen especially in the frequently mislocalized carboxysomes of the Δ*bolA* spherical mutant of *F. diplosiphon* (Figure 2.8 & Figure 5.10A). The functional importance of this is substantial, as cell shape was implicated in the determination of A in WT under GL (Figure 5.1C) and Δ*bolA* (Figure 5.1F).

Our understanding of A behavior in cyanobacteria has been expanded through the use of semi-wet *F. diplosiphon* discs. This technique allowed us to make CRC measurements comparable to the informative carbon assimilation vs intracellular CO_2 curves used in higher-order plants. Two major phases of CRCs were identified in *F. diplosiphon*: a C_i-uptake driven region and an upper region that includes A_{max} and is dependent on C_i-uptake, PSII rates, overall rate of linear electron flow, cell shape, and/or carboxysome morphology. These data highlight multiple responses by the CCM to environmental factors such as light intensity, light wavelength, and C_i availability. These responses were shown to influence A behavior, elucidating a functional role for fine-tuned CCM regulation on organismal fitness in cyanobacteria.

Ultimately, my studies draw strong correlation between the detection of changes in the external environment and tuning of cellular capacity to produce sustaining energy through photosynthesis. The multiple points of regulation of CCM, the investment in a regulatory system to drive fine-tuning, and the direct impact of these on organismal fitness provide strong evidence of the importance of regulating multiple points of photosynthesis to thrive in dynamic environments.

BIBLIOGRAPHY

BIBLIOGRAPHY

(1) Aggarwal, K. K., Saluja, D., and Sachar, R. C. (1993) Phosphorylation of rubisco in *Cicer rietinum*: Non-phosphoprotein nature of rubisco in *Nicotiana tabacun*. *Phytochemistry 34*, 329–335. doi: 10.1016/0031-9422(93)80004-C

(2) Aigner, H., Wilson, R. H., Bracher, A., Calisse, L., Bhat, J. Y., Hartl, F. U., and Hayer-Hartl, M. (2017) Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science 358*, 1272–1278. doi: 10.1126/science.aap9221

(3) Aldea, M., Garrido, T., Hernández-Chico, C., Vicente, M., and Kushner, S. R. (1989) Induction of a growth-phase-dependent promoter triggers transcription of *bolA*, an *Escherichia coli* morphogene. *EMBO J. 8*, 3923–3931.

(4) Alvey, R. M., Bezy, R. P., Frankenberg-Dinkel, N., and Kehoe, D. M. (2007) A light regulated OmpR-class promoter element co-ordinates light-harvesting protein and chromophore biosynthetic enzyme gene expression. *Mol. Microbiol. 64*, 319–332. doi: 10.1111/j.1365-2958.2007.05656.x

(5) Andrews, T. J., and Abel, K. M. (1981) Kinetics and subunit interactions of ribulose bisphosphate carboxylase-oxygenase from the cyanobacterium, *Synechococcus* sp. *J. Biol. Chem. 256*, 8445–8451.

(6) de Araujo, C., Arefeen, D., Tadesse, Y., Long, B. M., Price, G. D., Rowlett, R. S., Kimber, M. S., and Espie, G. S. (2014) Identification and characterization of a carboxysomal γ-carbonic anhydrase from the cyanobacterium *Nostoc* sp. PCC 7120. *Photosynth. Res.* doi: 10.1007/s11120-014-0018-4

(7) Aryal, U. K., Stöckel, J., Krovvidi, R. K., Gritsenko, M. A., Monroe, M. E., Moore, R. J., Koppenaal, D. W., Smith, R. D., Pakrasi, H. B., and Jacobs, J. M. (2011) Dynamic proteomic profiling of a unicellular cyanobacterium *Cyanothece* ATCC51142 across light-dark diurnal cycles. *BMC Syst. Biol. 5*, 194. doi: 10.1186/1752-0509-5-194

(8) Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (2014) Current Protocols in Molecular Biology. Greene Publishing and Wiley Interscience.

(9) Badger, M. R. (1980) Kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Anabaena variabilis*. *Arch. Biochem. Biophys. 201*, 247–254. doi: 10.1016/0003-9861(80)90509-3

(10) Badger, M. R., and Andrews, T. J. (1982) Photosynthesis and inorganic carbon usage by the marine cyanobacterium, *Synechococcus* sp. *Plant Physiol. 70*, 517–523.

(11) Badger, M. R., Andrews, T. J., Whitney, S. M., Ludwig, M., Yellowlees, D. C., Leggat, W., and Price, G. D. (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and

chloroplast-based CO2-concentrating mechanisms in algae. *Can. J. Bot. 76*, 1052–1071. doi: 10.1139/b98-074

(12) Badger, M. R., Hanson, D., and Price, G. D. (2002) Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Funct. Plant Biol. 29*, 161–173. doi: 10.1071/pp01213

(13) Badger, M. R., Palmqvist, K., and Yu, J.-W. (1994) Measurement of CO_2 and HCO_3^- fluxes in cyanobacteria and microalgae during steady-state photosynthesis. *Physiol. Plant. 90*, 529–536. doi: 10.1111/j.1399-3054.1994.tb08811.x

(14) Badger, M. R., and Price, G. (1992) The CO_2 concentrating mechanism in cyanobacteria and microalgae. *Physiol. Plant. 84*, 606–615. doi: 10.1034/j.1399-3054.1992.840416.x

(15) Ballaré, C. L., Scopel, A. L., and Sanchez, R. A. (1991) Photocontrol of stem elongation in plant neighbourhoods: effects of photon fluence rate under natural conditions of radiation. *Plant Cell Environ. 14*, 57–65. doi: 10.1111/j.1365-3040.1991.tb01371.x

(16) Bao, H., Melnicki, M. R., Pawlowski, E. G., Sutter, M., Agostoni, M., Lechno-Yossef, S., Cai, F., Montgomery, B. L., and Kerfeld, C. A. (2017) Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. *Nat. Plants 3*, 17089. doi: 10.1038/nplants.2017.89

(17) Barta, C., Carmo-Silva, A. E., and Salvucci, M. E. (2011) Rubisco activase activity assays, in *Photosynthesis Research Protocols* (Carpentier, R., Ed.), pp 375–382. Humana Press, Totowa, NJ. doi: 10.1007/978-1-60761-925-3_29

(18) Bassham, J. A., Krohne, S., and Lendzian, K. (1978) *In vivo* control mechanism of the carboxylation reaction, in *Photosynthetic Carbon Assimilation* (Siegelman, H. W., and Hind, G., Eds.), pp 77–93. Springer US, Boston, MA. doi: 10.1007/978-1-4684-8106-8_6

(19) Bennett, A., and Bogorad, L. (1973) Complementary chromatic adaptation in a filamentous blue-green alga. *J. Cell Biol. 58*, 419–435. doi: 10.1083/jcb.58.2.419

(20) Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., and Schwede, T. (2017) Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep. 7*. doi: 10.1038/s41598-017-09654-8

(21) Bezy, R. P., and Kehoe, D. M. (2010) Functional characterization of a cyanobacterial OmpR/PhoB class transcription factor binding site controlling light color responses. *J. Bacteriol. 192*, 5923–5933. doi: 10.1128/JB.00602-10

(22) Bhat, J. Y., Miličić, G., Thieulin-Pardo, G., Bracher, A., Maxwell, A., Ciniawsky, S., Mueller-Cajar, O., Engen, J. R., Hartl, F. U., Wendler, P., and Hayer-Hartl, M. (2017) Mechanism of enzyme repair by the AAA+ chaperone rubisco activase. *Mol. Cell 67*, 744-756.e6. doi: 10.1016/j.molcel.2017.07.004

(23) Bordowitz, J. R., and Montgomery, B. L. (2008) Photoregulation of cellular morphology during complementary chromatic adaptation requires sensor-kinase-class protein RcaE in *Fremyella diplosiphon*. *J. Bacteriol. 190*, 4069–4074. doi: 10.1128/JB.00018-08

(24) Bordowitz, J. R., Whitaker, M. J., and Montgomery, B. L. (2010) Independence and interdependence of the photoregulation of pigmentation and development in *Fremyella diplosiphon*. *Commun. Integr. Biol. 3*, 151–153.

(25) Bornefeld, T., and Simonis, W. (1974) Effects of light, temperature, pH, and inhibitors on the ATP level of the blue-green alga *Anacystic nidulans*. *Planta 115*, 309–318. doi: 10.1007/BF00388613

(26) Bottomley, P. J., and Stewart, W. D. P. (1976) ATP pools and transients in the blue-green alga, *Anabaena cylindrica*. *Arch. Microbiol. 108*, 249–258. doi: 10.1007/BF00454849

(27) Bracher, A., Whitney, S. M., Hartl, F. U., and Hayer-Hartl, M. (2017) Biogenesis and metabolic maintenance of rubisco. *Annu. Rev. Plant Biol. 68*, 29–60. doi: 10.1146/annurev-arplant-043015-111633

(28) Brooks, A., and Farquhar, G. D. (1985) Effect of temperature on the CO_2 / O_2 specificity of ribulose-1,5-bisphosphate carboxylase/oxygenase and the rate of respiration in the light. *Planta 165*, 397–406. doi: 10.1007/BF00392238

(29) Brown, M. R. W., and Kornberg, A. (2008) The long and short of it – polyphosphate, PPK and bacterial survival. *Trends Biochem. Sci. 33*, 284–290. doi: 10.1016/j.tibs.2008.04.005

(30) Burnap, R. L., Hagemann, M., and Kaplan, A. (2015) Regulation of CO_2 concentrating mechanism in cyanobacteria. *Life 5*, 348–371. doi: 10.3390/life5010348

(31) Burnap, R. L., Nambudiri, R., and Holland, S. (2013) Regulation of the carbon-concentrating mechanism in the cyanobacterium *Synechocystis* sp. PCC6803 in response to changing light intensity and inorganic carbon availability. *Photosynth. Res. 118*, 115–124. doi: 10.1007/s11120-013-9912-4

(32) Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., and Wittwer, C. T. (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem. 55*, 611–622. doi: 10.1373/clinchem.2008.112797

(33) Cai, F., Bernstein, S. L., Wilson, S. C., and Kerfeld, C. A. (2016) Production and characterization of synthetic carboxysome shells with incorporated luminal proteins. *Plant Physiol. 170*, 1868–1877. doi: 10.1104/pp.15.01822

(34) Cai, F., Menon, B. B., Cannon, G. C., Curry, K. J., Shively, J. M., and Heinhorst, S. (2009) The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier. *PLoS ONE 4*. doi: 10.1371/journal.pone.0007521

(35) Cai, F., Sutter, M., Bernstein, S. L., Kinney, J. N., and Kerfeld, C. A. (2015) Engineering bacterial microcompartment shells: Chimeric shell proteins and chimeric carboxysome shells. *ACS Synth. Biol. 4*, 444–453. doi: 10.1021/sb500226j

(36) Cai, F., Sutter, M., Cameron, J. C., Stanley, D. N., Kinney, J. N., and Kerfeld, C. A. (2013) The structure of CcmP, a tandem bacterial microcompartment domain protein from the β-carboxysome, forms a subcompartment within a microcompartment. *J. Biol. Chem. 288*, 16055–16063. doi: 10.1074/jbc.M113.456897

(37) Calvin, M., and Benson, A. A. (1948) The path of carbon in photosynthesis. *Science 107*, 476–480.

(38) Cameron, J. C., Wilson, S. C., Bernstein, S. L., and Kerfeld, C. A. (2013) Biogenesis of a bacterial organelle: The carboxysome assembly pathway. *Cell 155*, 1131–1140. doi: 10.1016/j.cell.2013.10.044

(39) Campbell, D. (1996) Complementary chromatic adaptation alters photosynthetic strategies in the cyanobacterium *Calothrix*. *Microbiology 142*, 1255–1263. doi: 10.1099/13500872-142-5-1255

(40) Casey, E. S., Kehoe, D. M., and Grossman, A. R. (1997) Suppression of mutants aberrant in light intensity responses of complementary chromatic adaptation. *J. Bacteriol. 179*, 4599–4606.

(41) Chaudhury, S., Berrondo, M., Weitzner, B. D., Muthu, P., Bergman, H., and Gray, J. J. (2011) Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PLoS ONE 6*. doi: 10.1371/journal.pone.0022477

(42) Cobley, J. G., Zerweck, E., Reyes, R., Mody, A., Seludo-Unson, J. R., Jaeger, H., Weerasuriya, S., and Navankasattusas, 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. doi: 10.1006/plas.1993.1037

(43) Cobley, J., Seneviratne, L., Drong, L., Thounaojam, M., Oda, J. F., and Carroll, J. (1999) Transposition of Tn5 derivatives in the chromatically adapting cyanobacterium, *Fremyella Diplosiphon*, in *The Phototrophic Prokaryotes* (Peschek, G. A., Löffelhardt, W., and Schmetterer, G., Eds.), pp 443–451. Springer US, Boston, MA. doi: 10.1007/978-1-4615-4827-0_52

(44) Cot, S. S.-W., So, A. K.-C., and Espie, G. S. (2008) A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J. Bacteriol. 190*, 936–945. doi: 10.1128/JB.01283-07

(45) Delwiche, C. F. (1999) Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat. 154*, S164–S177. doi: 10.1086/303291

(46) Dou, Z., Heinhorst, S., Williams, E. B., Murin, C. D., Shively, J. M., and Cannon, G. C. (2008) CO_2 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking

carbonic anhydrase suggest the shell acts as a diffusional barrier for CO_2. *J. Biol. Chem. 283*, 10377–10384. doi: 10.1074/jbc.M709285200

(47) Edgar, R. C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res. 32*, 1792–1797. doi: 10.1093/nar/gkh340

(48) Eisenhut, M., von Wobeser, E. A., Jonas, L., Schubert, H., Ibelings, B. W., Bauwe, H., Matthijs, H. C. P., and Hagemann, M. (2007) Long-term response toward inorganic carbon limitation in wild type and glycolate turnover mutants of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Physiol. 144*, 1946–1959. doi: 10.1104/pp.107.103341

(49) Elhai, J. (1993) Strong and regulated promoters in the cyanobacterium *Anabaena* PCC 7120. *FEMS Microbiol. Lett. 114*, 179–184. doi: 10.1111/j.1574-6968.1993.tb06570.x

(50) Elhai, J., Vepritskiy, A., Muro-Pastor, A. M., Flores, E., and Wolk, C. P. (1997) Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120. *J. Bacteriol. 179*, 1998–2005.

(51) Elhai, J., and Wolk, C. P. (1988) [83] Conjugal transfer of DNA to cyanobacteria, in *Methods in Enzymology*, pp 747–754. Academic Press. doi: 10.1016/0076-6879(88)67086-8

(52) Ellis, R. J. (1979) The most abundant protein in the world. *Trends Biochem. Sci. 4*, 241–244. doi: 10.1016/0968-0004(79)90212-3

(53) Erb, T. J., and Zarzycki, J. (2018) A short history of rubisco: the rise and fall (?) of nature's predominant CO_2 fixing enzyme. *Curr. Opin. Biotechnol. 49*, 100–107. doi: 10.1016/j.copbio.2017.07.017

(54) Espie, G. S., and Kimber, M. S. (2011) Carboxysomes: Cyanobacterial rubisco comes in small packages. *Photosynth. Res. 109*, 7–20. doi: 10.1007/s11120-011-9656-y

(55) Farquhar, G. D., von Caemmerer, S., and Berry, J. A. (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta 149*, 78–90.

(56) Faulkner, M., Rodriguez-Ramos, J., Dykes, G. F., Owen, S. V., Casella, S., Simpson, D. M., Beynon, R. J., and Liu, L.-N. (2017) Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes. *Nanoscale 9*, 10662–10673. doi: 10.1039/C7NR02524F

(57) Feiz, L., Williams-Carrier, R., Wostrikoff, K., Belcher, S., Barkan, A., and Stern, D. B. (2012) Ribulose-1,5-bis-phosphate carboxylase/oxygenase accumulation factor1 is required for holoenzyme assembly in Maize. *Plant Cell 24*, 3435–3446. doi: 10.1105/tpc.112.102012

(58) Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science 281*, 237–240. doi: 10.1126/science.281.5374.237

(59) Gantt, E., and Conti, S. F. (1969) Ultrastructure of blue-green algae. *J. Bacteriol. 97*, 1486–1493.

(60) Gaudana, S. B., Zarzycki, J., Moparthi, V. K., and Kerfeld, C. A. (2015) Bioinformatic analysis of the distribution of inorganic carbon transporters and prospective targets for bioengineering to increase C_i uptake by cyanobacteria. *Photosynth. Res. 126*, 99–109. doi: 10.1007/s11120-014-0059-8

(61) Gill, R. T., Katsoulakis, E., Schmitt, W., Taroncher-Oldenburg, G., Misra, J., and Stephanopoulos, G. (2002) Genome-wide dynamic transcriptional profiling of the light-to-dark transition in *Synechocystis* sp. strain PCC 6803. *J. Bacteriol. 184*, 3671–3681. doi: 10.1128/JB.184.13.3671-3681.2002

(62) Giovannoni, S. J., Turner, S., Olsen, G. J., Barns, S., Lane, D. J., and Pace, N. R. (1988) Evolutionary relationships among cyanobacteria and green chloroplasts. *J. Bacteriol. 170*, 3584–3592. doi: 10.1128/jb.170.8.3584-3592.1988

(63) Gomez-Garcia, M. R., Fazeli, F., Grote, A., Grossman, A. R., and Bhaya, D. (2013) Role of polyphosphate in thermophilic *Synechococcus* sp. from microbial mats. *J. Bacteriol. 195*, 3309–3319. doi: 10.1128/JB.00207-13

(64) Gonzalez-Esquer, C. R., Shubitowski, T. B., and Kerfeld, C. A. (2015) Streamlined construction of the cyanobacterial CO_2-fixing organelle via protein domain fusions for use in plant synthetic biology. *Plant Cell 27*, 2637–2644. doi: 10.1105/tpc.15.00329

(65) Gorelova, O. A., Baulina, O. I., Rasmussen, U., and Koksharova, O. A. (2013) The pleiotropic effects of ftn2 and ftn6 mutations in cyanobacterium *Synechococcus* sp. PCC 7942. *Protoplasma 250*, 931–942. doi: 10.1007/s00709-012-0479-2

(66) Graan, T., and Ort, D. R. (1986) Detection of oxygen-evolving Photosystem II centers inactive in plastoquinone reduction. *Biochim. Biophys. Acta Bioenerg. 852*, 320–330. doi: 10.1016/0005-2728(86)90238-0

(67) Grabsztunowicz, M., Koskela, M. M., and Mulo, P. (2017) Post-translational modifications in regulation of chloroplast function: Recent advances. *Front. Plant Sci. 8*. doi: 10.3389/fpls.2017.00240

(68) Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., and Baker, D. (2003) Protein–Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. *J. Mol. Biol. 331*, 281–299. doi: 10.1016/S0022-2836(03)00670-3

(69) Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol. 59*, 307–321. doi: 10.1093/sysbio/syq010

(70) Guitton, C., and Mache, R. (1987) Phosphorylation in vitro of the large subunit of the ribulose-1,5-bisphosphate carboxylase and of the glyceraldehyde-3-phosphate dehydrogenase. *Eur. J. Biochem. 166*, 249–254. doi: 10.1111/j.1432-1033.1987.tb13509.x

(71) Gutu, A., and Kehoe, D. M. (2012) Emerging perspectives on the mechanisms, regulation, and distribution of light color acclimation in cyanobacteria. *Mol. Plant 5*, 1–13. doi: 10.1093/mp/ssr054

(72) Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol. 177*, 4121–4130. doi: 10.1128/jb.177.14.4121-4130.1995

(73) Hasse, D., Larsson, A. M., and Andersson, I. (2015) Structure of *Arabidopsis thaliana* rubisco activase. *Acta Crystallogr. D Biol. Crystallogr. 71*, 800–808. doi: 10.1107/S1399004715001182

(74) He, Y.-Y., and Häder, D.-P. (2002) Involvement of reactive oxygen species in the UV-B damage to the cyanobacterium *Anabaena* sp. *J. Photochem. Photobiol. B 66*, 73–80. doi: 10.1016/S1011-1344(01)00278-0

(75) Heber, U., Takahama, U., Neimanis, S., and Shimizu-Takahama, M. (1982) Transport as the basis of the kok effect. Levels of some photosynthetic intermediates and activation of light-regulated enzymes during photosynthesis of chloroplasts and green leaf protoplasts. *Biochim. Biophys. Acta Bioenerg. 679*, 287–299. doi: 10.1016/0005-2728(82)90299-7

(76) Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikeuchi, M. (2001) DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *Plant Cell 13*, 793–806.

(77) Hosseinzadeh, P., Bhardwaj, G., Mulligan, V. K., Shortridge, M. D., Craven, T. W., Pardo-Avila, F., Rettie, S. A., Kim, D. E., Silva, D.-A., Ibrahim, Y. M., Webb, I. K., Cort, J. R., Adkins, J. N., Varani, G., and Baker, D. (2017) Comprehensive computational design of ordered peptide macrocycles. *Science 358*, 1461–1466. doi: 10.1126/science.aap7577

(78) Huang, J., Wang, J., and Xu, H. (2014) The circadian rhythms of photosynthesis, ATP content and cell division in *Microcystisaeruginosa* PCC7820. *Acta Physiol. Plant. 36*, 3315–3323. doi: 10.1007/s11738-014-1699-1

(79) Huang, L., McCluskey, M. P., Ni, H., and LaRossa, R. A. (2002) Global gene expression profiles of the cyanobacterium *Synechocystis* sp. strain PCC 6803 in response to irradiation with UV-B and white light. *J. Bacteriol. 184*, 6845–6858. doi: 10.1128/JB.184.24.6845-6858.2002

(80) Iancu, C. V., Ding, H. J., Morris, D. M., Dias, D. P., Gonzales, A. D., Martino, A., and Jensen, G. J. (2007) The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. *J. Mol. Biol. 372*, 764–773. doi: 10.1016/j.jmb.2007.06.059

(81) Ito, H., Mutsuda, M., Murayama, Y., Tomita, J., Hosokawa, N., Terauchi, K., Sugita, C., Sugita, M., Kondo, T., and Iwasaki, H. (2009) Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. 106*, 14168–14173. doi: 10.1073/pnas.0902587106

(82) Jensen, T., and Bowen, C. (1961) Organization of the centroplasm in *Nostoc pruniforme*. *Proc. Iowa Acad. Sci. 68*, 86–89.

(83) Jordan, D. B., and Chollet, R. (1983) Inhibition of ribulose bisphosphate carboxylase by substrate ribulose 1,5-bisphosphate. *J. Biol. Chem. 258*, 13752–13758.

(84) Kaiser, E., Morales, A., and Harbinson, J. (2018) Fluctuating light takes crop photosynthesis on a rollercoaster ride. *Plant Physiol. 176*, 977–989. doi: 10.1104/pp.17.01250

(85) Kakani, V. G., Surabhi, G. K., and Reddy, K. R. (2008) Photosynthesis and fluorescence responses of C_4 plant *Andropogon gerardii* acclimated to temperature and carbon dioxide. *Photosynthetica 46*, 420–430. doi: 10.1007/s11099-008-0074-0

(86) Kaplan, A., and Reinhold, L. (1999) CO_2 concentrating mechanisms in photosynthetic microorganisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol. 50*, 539–570. doi: 10.1146/annurev.arplant.50.1.539

(87) Karpova, T., and McNally, J. G. (2006) Detecting protein–protein interactions with CFP-YFP FRET by acceptor photobleaching. *Curr. Protoc. Cytom. 35*, 12.7.1-12.7.11. doi: 10.1002/0471142956.cy1207s35

(88) Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res. 33*, 511–518. doi: 10.1093/nar/gki198

(89) Kehoe, D. M., and Grossman, A. R. (1996) Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science 273*, 1409–1412. doi: 10.1126/science.273.5280.1409

(90) Kehoe, D. M., and Grossman, A. R. (1997) New classes of mutants in complementary chromatic adaptation provide evidence for a novel four-step phosphorelay system. *J. Bacteriol. 179*, 3914–3921.

(91) Kerfeld, C. A., and Melnicki, M. R. (2016) Assembly, function and evolution of cyanobacterial carboxysomes. *Curr. Opin. Plant Biol. 31*, 66–75. doi: 10.1016/j.pbi.2016.03.009

(92) Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M., and Yeates, T. O. (2005) Protein structures forming the shell of primitive bacterial organelles. *Science 309*, 936–938. doi: 10.1126/science.1113397

(93) Khan, S., Andralojc, P. J., Lea, P. J., and Parry, M. A. J. (1999) 2′-Carboxy-D-arabitinol 1-phosphate protects ribulose 1,5-bisphosphate carboxylase/oxygenase against proteolytic breakdown. *Eur. J. Biochem. 266*, 840–847. doi: 10.1046/j.1432-1327.1999.00913.x

(94) Kinney, J. N., Salmeen, A., Cai, F., and Kerfeld, C. A. (2012) Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. *J. Biol. Chem. 287*, 17729–17736. doi: 10.1074/jbc.M112.355305

(95) Knight, S., Andersson, I., and Brändén, C.-I. (1990) Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2·4 Å resolution: Subunit interactions and active site. *J. Mol. Biol. 215*, 113–160. doi: 10.1016/S0022-2836(05)80100-7

(96) Kobayashi, Y., Inoue, Y., Shibata, K., and Heber, U. (1979) Control of electron flow in intact chloroplasts by the intrathylakoid pH, not by the phosphorylation potential. *Planta 146*, 481–486. doi: 10.1007/BF00380864

(97) Koksharova, O. A., and Wolk, C. P. (2002) A novel gene that bears a DnaJ motif influences cyanobacterial cell division. *J. Bacteriol. 184*, 5524–5528. doi: 10.1128/JB.184.19.5524-5528.2002

(98) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 774–797. doi: 10.1016/j.jmb.2007.05.022

(99) Kuriata, A., Chakraborty, M., Nathan Henderson, J., Hazra, S., J Serban, A., V T Pham, T., Levitus, M., and Maria Wachter, R. (2014) ATP and magnesium promote cotton short-form ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) activase hexamer formation at low micromolar concentrations. *Biochemistry 53*. doi: 10.1021/bi500968h

(100) Laskowski, R. A. (2017) The ProFunc function prediction server. *Methods Mol. Biol. Clifton NJ 1611*, 75–95. doi: 10.1007/978-1-4939-7015-5_7

(101) Lechno-Yossef, S., Rohnke, B. A., Belza, A. C. O., Melnicki, M. R., Montgomery, B. L., and Kerfeld, C. A. (2019) Cyanobacterial carboxysomes contain an unique rubisco-activase-like protein. *New Phytol. Early View*. doi: 10.1111/nph.16195

(102) Li, L. A., and Tabita, F. R. (1997) Maximum activity of recombinant ribulose 1,5-bisphosphate carboxylase/oxygenase of *Anabaena* sp. strain CA requires the product of the rbcX gene. *J. Bacteriol. 179*, 3793–3796.

(103) Li, L. A., Zianni, M. R., and Tabita, F. R. (1999) Inactivation of the monocistronic *rca* gene in *Anabaena variabilis* suggests a physiological ribulose bisphosphate carboxylase/oxygenase activase-like function in heterocystous cyanobacteria. *Plant Mol. Biol. 40*, 467–478. doi: 10.1023/a:1006251808625

(104) Li, L., Alvey, R. M., Bezy, R. P., and Kehoe, D. M. (2008) Inverse transcriptional activities during complementary chromatic adaptation are controlled by the response regulator RcaC binding to red and green light-responsive promoters. *Mol. Microbiol. 68*, 286–297. doi: 10.1111/j.1365-2958.2008.06151.x

(105) Li, L., and Kehoe, D. M. (2005) *In vivo* analysis of the roles of conserved aspartate and histidine residues within a complex response regulator. *Mol. Microbiol. 55*, 1538–1552. doi: 10.1111/j.1365-2958.2005.04491.x

(106) Li, L., and Kehoe, D. M. (2008) Abundance changes of the response regulator RcaC require specific aspartate and histidine residues and are necessary for normal light color responsiveness. *J. Bacteriol. 190*, 7241–7250. doi: 10.1128/JB.00762-08

(107) Li, L.-A., Janet, L., Gibson, and Robert Tabita, F. (1993) The rubisco activase (*rca*) gene is located downstream from *rbcS* in *Anabaena* sp. strain CA and is detected in other *Anabaena*/*Nostoc* strains. *Plant Mol. Biol. 21*, 753–764. doi: 10.1007/BF00027109

(108) Liang, Y., Urano, D., Liao, K.-L., Hedrick, T. L., Gao, Y., and Jones, A. M. (2017) A nondestructive method to estimate the chlorophyll content of *Arabidopsis* seedlings. *Plant Methods 13*, 26. doi: 10.1186/s13007-017-0174-6

(109) Liberton, M., Austin, J. R., Berg, R. H., and Pakrasi, H. B. (2011) Unique thylakoid membrane architecture of a unicellular N_2-fixing cyanobacterium revealed by electron tomography. *Plant Physiol. 155*, 1656–1666. doi: 10.1104/pp.110.165332

(110) Liu, H., Landry, M. R., Vaulot, D., and Campbell, L. (1999) *Prochlorococcus* growth rates in the central equatorial Pacific: An application of the *ƒ*max approach. *J. Geophys. Res. Oceans 104*, 3391–3399. doi: 10.1029/1998JC900011

(111) Lohrig, K., Müller, B., Davydova, J., Leister, D., and Wolters, D. A. (2009) Phosphorylation site mapping of soluble proteins: Bioinformatical filtering reveals potential plastidic phosphoproteins in *Arabidopsis* thaliana. *Planta 229*, 1123–1134. doi: 10.1007/s00425-009-0901-y

(112) Long, B. M., Badger, M. R., Whitney, S. M., and Price, G. D. (2007) Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple rubisco complexes with carboxysomal proteins CcmM and CcaA. *J. Biol. Chem. 282*, 29323–29335. doi: 10.1074/jbc.M703896200

(113) Long, B. M., Rae, B. D., Badger, M. R., and Price, G. D. (2011) Over-expression of the β-carboxysomal CcmM protein in *Synechococcus* PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. *Photosynth. Res. 109*, 33–45. doi: 10.1007/s11120-011-9659-8

(114) Long, B. M., Tucker, L., Badger, M. R., and Price, G. D. (2010) Functional cyanobacterial β-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. *Plant Physiol. 153*, 285–293. doi: 10.1104/pp.110.154948

(115) Long, S. P., and Bernacchi, C. J. (2003) Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *J. Exp. Bot. 54*, 2393–2401.

(116) Ludwig, M., Sultemeyer, D., and Price, G. D. (2000) Isolation of *ccmKLMN* genes from the marine cyanobacterium, *Synechococcus* sp PCC7002 (*Cyanophyceae*) and evidence that CcmM is essential for carboxysome assembly. *J. Phycol. 36*, 1109–1119. doi: 10.1046/j.1529-8817.2000.00028.x

(117) van Lun, M., van der Spoel, D., and Andersson, I. (2011) Subunit interface dynamics in hexadecameric rubisco. *J. Mol. Biol. 411*, 1083–1098. doi: 10.1016/j.jmb.2011.06.052

(118) MacCready, J. S., Hakim, P., Young, E. J., Hu, L., Liu, J., Osteryoung, K. W., Vecchiarelli, A. G., and Ducat, D. C. (2018) Protein gradients on the nucleoid position the carbon-fixing organelles of cyanobacteria. *eLife 7*, e39723. doi: doi.org/10.7554/eLife.39723

(119) Mackinder, L. C. M., Chen, C., Leib, R. D., Patena, W., Blum, S. R., Rodman, M., Ramundo, S., Adams, C. M., and Jonikas, M. C. (2017) A spatial interactome reveals the protein organization of the algal CO_2-concentrating mechanism. *Cell 171*, 133-147.e14. doi: 10.1016/j.cell.2017.08.044

(120) Mahinthichaichan, P., Morris, D. M., Wang, Y., Jensen, G. J., and Tajkhorshid, E. (2018) Selective permeability of carboxysome shell pores to anionic molecules. *bioRxiv*. doi: 10.1101/367714

(121) Mangan, N. M., and Brenner, M. P. (2014) Systems analysis of the CO_2 concentrating mechanism in cyanobacteria. *eLife* (Milo, R., Ed.) *3*, e02043. doi: 10.7554/eLife.02043

(122) de Marsac, N. T., and Houmard, J. (1988) [34] Complementary chromatic adaptation: Physiological conditions and action spectra, in *Methods in Enzymology*, pp 318–328. Academic Press. doi: 10.1016/0076-6879(88)67037-6

(123) Mcginn, P. J., Price, G. D., and Badger, M. R. (2004) High light enhances the expression of low-CO_2-inducible transcripts involved in the CO_2-concentrating mechanism in *Synechocystis* sp. PCC6803. *Plant Cell Environ. 27*, 615–626. doi: 10.1111/j.1365-3040.2004.01175.x

(124) McGinn, P. J., Price, G. D., Maleszka, R., and Badger, M. R. (2003) Inorganic carbon limitation and light control the expression of transcripts related to the CO_2-concentrating mechanism in the cyanobacterium *Synechocystis* sp. strain PCC6803. *Plant Physiol. 132*, 218–229. doi: 10.1104/pp.019349

(125) McGurn, L. D., Moazami-Goudarzi, M., White, S. A., Suwal, T., Brar, B., Tang, J. Q., Espie, G. S., and Kimber, M. S. (2016) The structure, kinetics and interactions of the β-carboxysomal β-carbonic anhydrase, CcaA. *Biochem. J. 473*, 4559–4572. doi: 10.1042/BCJ20160773

(126) McKay, R. M. L., and Gibbs, S. P. (1991) Composition and function of pyrenoids: Cytochemical and immunocytochemical approaches. *Can. J. Bot. 69*, 1040–1052. doi: 10.1139/b91-134

(127) McKay, R. M. L., Gibbs, S. P., and Espie, G. S. (1993) Effect of dissolved inorganic carbon on the expression of carboxysomes, localization of Rubisco and the mode of inorganic carbon transport in cells of the cyanobacterium Synechococcus UTEX 625. *Arch. Microbiol. 159*, 21–29. doi: 10.1007/BF00244259

(128) Menon, B. B., Heinhorst, S., Shively, J. M., and Cannon, G. C. (2010) The carboxysome shell is permeable to protons. *J. Bacteriol. 192*, 5881–5886. doi: 10.1128/JB.00903-10

(129) Montgomery, B. L. (2016) Mechanisms and fitness implications of photomorphogenesis during chromatic acclimation in cyanobacteria. *J. Exp. Bot. 67*, 4079–4090. doi: 10.1093/jxb/erw206

(130) Montgomery, B. L., Lechno-Yossef, S., and Kerfeld, C. A. (2016) Interrelated modules in cyanobacterial photosynthesis: the carbon-concentrating mechanism, photorespiration, and light perception. *J. Exp. Bot. 67*, 2931–2940. doi: 10.1093/jxb/erw162

(131) Mueller-Cajar, O. (2017) The diverse AAA+ machines that repair inhibited rubisco active sites. *Front. Mol. Biosci. 4*. doi: 10.3389/fmolb.2017.00031

(132) Mueller-Cajar, O., Stotz, M., Wendler, P., Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Structure and function of the AAA+ protein CbbX, a red-type rubisco activase. *Nature 479*, 194–199. doi: 10.1038/nature10568

(133) Mulo, P., Laakso, S., Mäenpää, P., and Aro, E.-M. (1998) Stepwise photoinhibition of Photosystem II. *Plant Physiol. 117*, 483–490.

(134) Murata, K., Hagiwara, S., Kimori, Y., and Kaneko, Y. (2016) Ultrastructure of compacted DNA in cyanobacteria by high-voltage cryo-electron tomography. *Sci. Rep. 6*, 34934. doi: 10.1038/srep34934

(135) Nagarajan, R., and Gill, K. S. (2018) Evolution of rubisco activase gene in plants. *Plant Mol. Biol. 96*, 69–87. doi: 10.1007/s11103-017-0680-y

(136) Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res. 9*, 27–43. doi: 10.1101/gr.9.1.27

(137) Niederhuber, M. J., Lambert, T. J., Yapp, C., Silver, P. A., and Polka, J. K. (2017) Superresolution microscopy of the β-carboxysome reveals a homogeneous matrix. *Mol. Biol. Cell 28*, 2734–2745. doi: 10.1091/mbc.E17-01-0069

(138) Nierzwicki-Bauer, S. A., Balkwill, D. L., and Stevens, S. E., Jr. (1983) Three-dimensional ultrastructure of a unicellular cyanobacterium. *J. Cell Biol. 97*, 713–722.

(139) Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998) Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol. 64*, 1694–1699.

(140) Nitta, K., Nagayama, K., Danev, R., and Kaneko, Y. (2009) Visualization of BrdU-labelled DNA in cyanobacterial cells by Hilbert differential contrast transmission electron microscopy. *J. Microsc. 234*, 118–123. doi: 10.1111/j.1365-2818.2009.03162.x

(141) Oakley, C. A., Hopkinson, B. M., and Schmidt, G. W. (2012) A modular system for the measurement of CO_2 and O_2 gas flux and photosynthetic electron transport in microalgae. *Limnol. Oceanogr. Methods 10*, 968–977. doi: 10.4319/lom.2012.10.968

(142) Ogawa, T., Amichay, D., and Gurevitz, M. (1994) Isolation and characterization of the *ccmM* gene required by the cyanobacterium *Synechocystis* PCC6803 for inorganic carbon utilization. *Photosynth. Res. 39*, 183–190. doi: 10.1007/BF00029385

(143) Omata, T., Gohta, S., Takahashi, Y., Harano, Y., and Maeda, S. (2001) Involvement of a CbbR homolog in low CO_2-induced activation of the bicarbonate transporter operon in cyanobacteria. *J. Bacteriol. 183*, 1891–1898. doi: 10.1128/JB.183.6.1891-1898.2001

(144) Omata, T., Price, G. D., Badger, M. R., Okamura, M., Gohta, S., and Ogawa, T. (1999) Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Proc. Natl. Acad. Sci. 96*, 13571–13576. doi: 10.1073/pnas.96.23.13571

(145) Parry, M. A. J., Keys, A. J., Madgwick, P. J., Carmo-Silva, A. E., and Andralojc, P. J. (2008) Rubisco regulation: A role for inhibitors. *J. Exp. Bot. 59*, 1569–1580.

(146) Pattanaik, B., Busch, A. W. U., Hu, P., Chen, J., and Montgomery, B. L. (2014) Responses to iron limitation are impacted by light quality and regulated by RcaE in the chromatically acclimating cyanobacterium *Fremyella diplosiphon*. *Microbiology, 160*, 992–1005. doi: 10.1099/mic.0.075192-0

(147) Pearce, F. G. (2006) Catalytic by-product formation and ligand binding by ribulose bisphosphate carboxylases from different phylogenies. *Biochem. J. 399*, 525–534. doi: 10.1042/BJ20060430

(148) Peña, K. L., Castel, S. E., Araujo, C. de, Espie, G. S., and Kimber, M. S. (2010) Structural basis of the oxidative activation of the carboxysomal γ-carbonic anhydrase, CcmM. *Proc. Natl. Acad. Sci. 107*, 2455–2460. doi: 10.1073/pnas.0910866107

(149) Peterson-Forbrook, D. S., Hilton, M. T., Tichacek, L., Henderson, J. N., Bui, H. Q., and Wachter, R. M. (2017) Nucleotide dependence of subunit rearrangements in short-form rubisco activase from spinach. *Biochemistry 56*, 4906–4921. doi: 10.1021/acs.biochem.7b00574

(150) Pfeffer, M., and Peisker, M. (1998) CO_2 gas exchange and phosphoenolpyruvate carboxylase activity in leaves of *Zea mays* L. *Photosynth. Res. 58*, 11. doi: 10.1023/A:1006188705423

(151) Pollock, S. V., Colombo, S. L., Prout, D. L., Godfrey, A. C., and Moroney, J. V. (2003) Rubisco activase is required for optimal photosynthesis in the green alga *Chlamydomonas reinhardtii* in a low-CO_2 atmosphere. *Plant Physiol. 133*, 1854–1861. doi: 10.1104/pp.103.032078

(152) Portis, A. R., Salvucci, M. E., and Ogren, W. L. (1986) Activation of ribulosebisphosphate carboxylase/oxygenase at physiological CO_2 and ribulosebisphosphate concentrations by rubisco activase. *Plant Physiol. 82*, 967–971.

(153) Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018) HMMER web server: 2018 update. *Nucleic Acids Res. 46*, W200–W204. doi: 10.1093/nar/gky448

(154) Price, G. D., and Badger, M. R. (1989) Isolation and characterization of high CO_2-requiring-mutants of the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol. 91*, 514–525.

(155) Price, G. D., Badger, M. R., Woodger, F. J., and Long, B. M. (2008) Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J. Exp. Bot. 59*, 1441–1461. doi: 10.1093/jxb/erm112

(156) Price, G. D., Howitt, S. M., Harrison, K., and Badger, M. R. (1993) Analysis of a genomic DNA region from the cyanobacterium *Synechococcus* sp. strain PCC7942 involved in carboxysome assembly and function. *J. Bacteriol. 175*, 2871–2879. doi: 10.1128/jb.175.10.2871-2879.1993

(157) Price, G. D., Woodger, F. J., Badger, M. R., Howitt, S. M., and Tucker, L. (2004) Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. *Proc. Natl. Acad. Sci. 101*, 18228–18233. doi: 10.1073/pnas.0405211101

(158) R Core Team. (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

(159) Rae, B. D., Long, B. M., Badger, M. R., and Price, G. D. (2012) Structural determinants of the outer shell of β-Carboxysomes in *Synechococcus elongatus* PCC 7942: Roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS ONE* (Lei, B., Ed.) *7*, e43871. doi: 10.1371/journal.pone.0043871

(160) Rae, B. D., Long, B. M., Badger, M. R., and Price, G. D. (2013) Functions, compositions, and evolution of the two types of carboxysomes: Polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. *Microbiol. Mol. Biol. Rev. 77*, 357–379. doi: 10.1128/MMBR.00061-12

(161) Rausenberger, J., Hussong, A., Kircher, S., Kirchenbauer, D., Timmer, J., Nagy, F., Schäfer, E., and Fleck, C. (2010) An integrative model for phytochrome B mediated photomorphogenesis: From protein dynamics to physiology. *PLoS ONE 5*. doi: 10.1371/journal.pone.0010721

(162) Read, B. A., and Tabita, F. R. (1992) A hybrid ribulose bisphosphate carboxylase/oxygenase enzyme exhibiting a substantial increase in substrate specificity factor. *Biochemistry 31*, 5553–5560. doi: 10.1021/bi00139a018

(163) Read, B. A., and Tabita, F. R. (1994) High substrate specificity factor ribulose bisphosphate carboxylase/oxygenase from eukaryotic marine algae and properties of recombinant cyanobacterial rubisco containing “algal” residue modifications. *Arch. Biochem. Biophys. 312*, 210–218. doi: 10.1006/abbi.1994.1301

(164) Reed, J. W., Nagatani, A., Elich, T. D., Fagan, M., and Chory, J. (1994) Phytochrome A and phytochrome B have overlapping but distinct functions in *Arabidopsis* development. *Plant Physiol. 104*, 1139–1149.

(165) Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., and Stanier, R. Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology, 111*, 1–61. doi: 10.1099/00221287-111-1-1

(166) Roach, T., and Krieger-Liszkay, A. K. (2014) Regulation of photosynthetic electron transport and photoinhibition. *Curr. Protein Pept. Sci. 15*, 351–362. doi: 10.2174/1389203715666140327105143

(167) Robinson, S. P., and Portis, A. R. (1989) Adenosine triphosphate hydrolysis by purified rubisco activase. *Arch. Biochem. Biophys. 268*, 93–99. doi: 10.1016/0003-9861(89)90568-7

(168) Rockwell, N. C., and Lagarias, J. C. (2010) A brief history of phytochromes. *Chemphyschem 11*, 1172–1180. doi: 10.1002/cphc.200900894

(169) Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green cyanobacteriochromes: Sensors of color and power. *Biochemistry 51*, 9667–9677. doi: 10.1021/bi3013565

(170) Rohnke, B. A., Singh, S. P., Pattanaik, B., and Montgomery, B. L. (2018) RcaE-dependent regulation of carboxysome structural proteins has a central role in environmental determination of carboxysome morphology and abundance in *Fremyella diplosiphon*. *mSphere 3*, e00617-17. doi: 10.1128/mSphere.00617-17

(171) Ryan, P., Forrester, T. J. B., Wroblewski, C., Kenney, T. M. G., Kitova, E. N., Klassen, J. S., and Kimber, M. S. (2018) The small RbcS-like domains of the β-carboxysome structural protein, CcmM, bind rubisco at a site distinct from that binding the RbcS subunit. *J. Biol. Chem. 294*, 2593–2603. doi: 10.1074/jbc.RA118.006330

(172) Salvucci, M. E., and Ogren, W. L. (1996) The mechanism of rubisco activase: Insights from studies of the properties and structure of the enzyme. *Photosynth. Res. 47*, 1–11. doi: 10.1007/BF00017748

(173) Santos, J. M., Freire, P., Vicente, M., and Arraiano, C. M. (1999) The stationary-phase morphogene *bolA* from *Escherichia coli* is induced by stress during early stages of growth. *Mol. Microbiol. 32*, 789–798. doi: 10.1046/j.1365-2958.1999.01397.x

(174) Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. A. (2010) Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science 327*, 1258–1261. doi: 10.1126/science.1186090

(175) Schaefer, M. R., Chiang, G. G., Cobley, J. G., and Grossman, A. R. (1993) Plasmids from two morphologically distinct cyanobacterial strains share a novel replication origin. *J. Bacteriol. 175*, 5701–5705.

(176) Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods 9*, 671–675.

(177) Schopf, J. W. (2012) The fossil record of cyanobacteria, in *Ecology of cyanobacteria II: Their diversity in space and time* (Whitton, B. A., Ed.), pp 15–36. Springer, New York.

(178) Schuster-Böckler, B., Schultz, J., and Rahmann, S. (2004) HMM Logos for visualization of protein families. *BMC Bioinformatics 5*, 7. doi: 10.1186/1471-2105-5-7

(179) Schwarz, D., Nodop, A., Hüge, J., Purfürst, S., Forchhammer, K., Michel, K.-P., Bauwe, H., Kopka, J., and Hagemann, M. (2011) Metabolic and transcriptomic phenotyping of inorganic carbon acclimation in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant Physiol. 155*, 1640–55.

(180) Seib, L. O., and Kehoe, D. M. (2002) A turquoise mutant genetically separates expression of genes encoding phycoerythrin and its associated linker peptides. *J. Bacteriol. 184*, 962–970. doi: 10.1128/jb.184.4.962-970.2002

(181) Seki, Y., Nitta, K., and Kaneko, Y. (2014) Observation of polyphosphate bodies and DNA during the cell division cycle of *Synechococcus elongatus* PCC 7942. *Plant Biol. 16*, 258–263. doi: 10.1111/plb.12008

(182) Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem. 68*, 850–858. doi: 10.1021/ac950914h

(183) Shibata, M., Katoh, H., Sonoda, M., Ohkawa, H., Shimoyama, M., Fukuzawa, H., Kaplan, A., and Ogawa, T. (2002) Genes essential to sodium-dependent bicarbonate transport in cyanobacteria FUNCTION AND PHYLOGENETIC ANALYSIS. *J. Biol. Chem. 277*, 18658–18664. doi: 10.1074/jbc.M112468200

(184) Shibata, M., Ohkawa, H., Kaneko, T., Fukuzawa, H., Tabata, S., Kaplan, A., and Ogawa, T. (2001) Distinct constitutive and low-CO_2-induced CO_2 uptake systems in cyanobacteria: Genes involved and their phylogenetic relationship with homologous genes in other organisms. *Proc. Natl. Acad. Sci. 98*, 11789–11794. doi: 10.1073/pnas.191258298

(185) Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Calteau, A., Cai, F., Marsac, N. T. de, Rippka, R., Herdman, M., Sivonen, K., Coursin, T., Laurent, T., Goodwin, L., Nolan, M., Davenport, K. W., Han, C. S., Rubin, E. M., Eisen, J. A., Woyke, T., Gugger, M., and Kerfeld, C. A. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. 110*, 1053–1058. doi: 10.1073/pnas.1217107110

(186) Singh, shailendra pratap, and L Montgomery, B. (2012) Reactive oxygen species are involved in the morphology-determining mechanism of *Fremyella diplosiphon* cells during complementary chromatic adaptation. *Microbiology 158*, 2235–45. doi: 10.1099/mic.0.060475-0

(187) Singh, S. P., and Montgomery, B. L. (2011) Temporal responses of wild-type pigmentation and RcaE-deficient strains of *Fremyella diplosiphon* during light transitions. *Commun. Integr. Biol. 4*, 503–510. doi: 10.4161/cib.16788

(188) Singh, S. P., and Montgomery, B. L. (2013) Distinct salt-dependent effects impair *Fremyella diplosiphon* pigmentation and cellular shape. *Plant Signal. Behav. 8*, e24713. doi: 10.4161/psb.24713

(189) Singh, S. P., and Montgomery, B. L. (2014) Morphogenes *bolA* and *mreB* mediate the photoregulation of cellular morphology during complementary chromatic acclimation in *Fremyella diplosiphon*. *Mol. Microbiol. 93*, 167–182. doi: 10.1111/mmi.12649

(190) Singh, S. P., and Montgomery, B. L. (2015) Regulation of BolA abundance mediates morphogenesis in *Fremyella diplosiphon*. *Front. Microbiol. 6*. doi: 10.3389/fmicb.2015.01215

(191) Somerville, C. R., Portis, A. R., and Ogren, W. L. (1982) A mutant of *Arabidopsis thaliana* which lacks activation of RuBP carboxylase *in vivo*. *Plant Physiol. 70*, 381–387.

(192) Sommer, M., Cai, F., Melnicki, M., and Kerfeld, C. A. (2017) β-Carboxysome bioinformatics: identification and evolution of new bacterial microcompartment protein gene classes and core locus constraints. *J. Exp. Bot. 68*, 3841–3855. doi: 10.1093/jxb/erx115

(193) Sommer, M., Sutter, M., Gupta, S., Kirst, H., Turmo, A., Lechno-Yossef, S., Burton, R. L., Saechao, C., Sloan, N. B., Cheng, X., Chan, L.-J. G., Petzold, C. J., Fuentes-Cabrera, M., Ralston, C. Y., and Kerfeld, C. A. (2019) Heterohexamers formed by CcmK3 and CcmK4 increase the complexity of beta carboxysome shells. *Plant Physiol. 179*, 156–167. doi: 10.1104/pp.18.01190

(194) Spreitzer, R. J., and Salvucci, M. E. (2002) Rubisco: Structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol. 53*, 449–475. doi: 10.1146/annurev.arplant.53.100301.135233

(195) Stamatakis, A. (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics 30*, 1312–1313. doi: 10.1093/bioinformatics/btu033

(196) Stotz, M., Mueller-cajar, O., Ciniawsky, S., Wendler, P., Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Structure of green-type rubisco activase from tobacco. *Nat. Struct. Mol. Biol. 18*, 1366–70. doi: 10.1038/nsmb.2171

(197) Stowe-Evans, E. L., Ford, J., and Kehoe, D. M. (2004) Genomic DNA microarray analysis: Identification of new genes regulated by light color in the cyanobacterium *Fremyella diplosiphon*. *J. Bacteriol. 186*, 4338–4349. doi: 10.1128/JB.186.13.4338-4349.2004

(198) Sültemeyer, D., Klughammer, B., Ludwig, M., Badger, M. R., and Price, G. D. (1997) Random insertional mutagenesis used in the generation of mutants of the marine cyanobacterium *Synechococcus* sp. Strain PCC7002 with an impaired CO_2 concentrating mechanism. *Funct. Plant Biol. 24*, 317–327. doi: 10.1071/pp96124

(199) Sun, Y., Casella, S., Fang, Y., Huang, F., Faulkner, M., Barrett, S., and Liu, L.-N. (2016) Light modulates the biosynthesis and organization of cyanobacterial carbon fixation machinery through photosynthetic electron flow. *Plant Physiol. 171*, 530–541. doi: 10.1104/pp.16.00107

(200) Sun, Y., Wollman, A. J. M., Huang, F., Leake, M. C., and Liu, L.-N. (2019) Single-organelle quantification reveals stoichiometric and structural variability of carboxysomes dependent on the environment. *Plant Cell 31*, 1648–1664. doi: 10.1105/tpc.18.00787

(201) Sutter, M., Greber, B., Aussignargues, C., and Kerfeld, C. A. (2017) Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. *Science 356*, 1293–1297. doi: 10.1126/science.aan3289

(202) Sutter, M., Roberts, E. W., Gonzalez, R. C., Bates, C., Dawoud, S., Landry, K., Cannon, G. C., Heinhorst, S., and Kerfeld, C. A. (2015) Structural characterization of a newly identified component of α-carboxysomes: The AAA+ domain protein CsoCbbQ. *Sci. Rep. 5*, 16243. doi: 10.1038/srep16243

(203) Sutter, M., Wilson, S. C., Deutsch, S., and Kerfeld, C. A. (2013) Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species. *Photosynth. Res. 118*, 9–16. doi: 10.1007/s11120-013-9909-z

(204) Tabita, F. R. (1999) Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. *Photosynth. Res. 60*, 1–28. doi: 10.1023/A:1006211417981

(205) Tabita, F. R., and Colletti, C. (1979) Carbon dioxide assimilation in cyanobacteria: Regulation of ribulose, 1,5-bisphosphate carboxylase. *J. Bacteriol. 140*, 452–458.

(206) Tanaka, S., Kerfeld, C. A., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C., and Yeates, T. O. (2008) Atomic-level models of the bacterial carboxysome shell. *Science 319*, 1083–1086. doi: 10.1126/science.1151458

(207) Tcherkez, G. G. B., Farquhar, G. D., and Andrews, T. J. (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci. 103*, 7246–7251. doi: 10.1073/pnas.0600605103

(208) Tchernov, D., Hassidim, M., Luz, B., Sukenik, A., Reinhold, L., and Kaplan, A. (1997) Sustained net CO_2 evolution during photosynthesis by marine microorganism. *Curr. Biol. 7*, 723–728. doi: 10.1016/S0960-9822(06)00330-7

(209) Terauchi, K., Montgomery, B. L., Grossman, A. R., Lagarias, J. C., and Kehoe, D. M. (2004) RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol. Microbiol. 51*, 567–577. doi: 10.1046/j.1365-2958.2003.03853.x

(210) Turmo, A., Gonzalez-Esquer, C. R., and Kerfeld, C. A. (2017) Carboxysomes: Metabolic modules for CO_2 fixation. *FEMS Microbiol. Lett. 364*. doi: 10.1093/femsle/fnx176

(211) Ulijasz, A. T., and Vierstra, R. D. (2011) Phytochrome structure and photochemistry: Recent advances toward a complete molecular picture. *Curr. Opin. Plant Biol. 14*, 498–506. doi: 10.1016/j.pbi.2011.06.002

(212) Wang, C., Schueler-Furman, O., and Baker, D. (2005) Improved side-chain modeling for protein–protein docking. *Protein Sci. Publ. Protein Soc. 14*, 1328–1339. doi: 10.1110/ps.041222905

(213) Wang, H., Yan, X., Aigner, H., Bracher, A., Nguyen, N. D., Hee, W. Y., Long, B. M., Price, G. D., Hartl, F. U., and Hayer-Hartl, M. (2019) Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature 566*, 131. doi: 10.1038/s41586-019-0880-5

(214) Wang, H.-L., Postier, B. L., and Burnap, R. L. (2004) Alterations in global patterns of gene expression in *Synechocystis* sp. PCC 6803 in response to inorganic carbon limitation and the inactivation of ndhR, a LysR family regulator. *J. Biol. Chem. 279*, 5739–5751. doi: 10.1074/jbc.M311336200

(215) Wang, Q., Serban, A. J., Wachter, R. M., and Moerner, W. E. (2018) Single-molecule diffusometry reveals the nucleotide-dependent oligomerization pathways of *Nicotiana tabacum* rubisco activase. *J. Chem. Phys. 148*, 123319. doi: 10.1063/1.5005930

(216) Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res. gky427*. doi: 10.1093/nar/gkg520

(217) Watson, G. M., and Tabita, F. R. (1996) Regulation, unique gene organization, and unusual primary structure of carbon fixation genes from a marine phycoerythrin-containing cyanobacterium. *Plant Mol. Biol. 32*, 1103–1115. doi: 10.1007/bf00041394

(218) Wheeler, T. J., Clements, J., and Finn, R. D. (2014) Skylign: A tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics 15*, 7. doi: 10.1186/1471-2105-15-7

(219) Whitehead, L., Long, B. M., Price, G. D., and Badger, M. R. (2014) Comparing the *in vivo* function of α-carboxysomes and β-carboxysomes in two model cyanobacteria. *Plant Physiol. 165*, 398–411. doi: 10.1104/pp.114.237941

(220) Whitton, B. A., and Potts, M. (2012) Introduction to the cyanobacteria, in *Ecology of cyanobacteria II: Their diversity in space and time* (Whitton, B. A., Ed.), pp 1–13. Springer, New York.

(221) Zacarias, L., and Reid, M. S. (1990) Role of growth regulators in the senescence of *Arabidopsis thaliana* leaves. *Physiol. Plant. 80*, 549–554. doi: 10.1111/j.1399-3054.1990.tb05677.x

(222) Zarzycki, J., Axen, S. D., Kinney, J. N., and Kerfeld, C. A. (2013) Cyanobacterial-based approaches to improving photosynthesis in plants. *J. Exp. Bot. 64*, 787–798. doi: 10.1093/jxb/ers294

(223) Zhang, N., Kallis, R. P., Ewy, R. G., and Portis, A. R. (2002) Light modulation of rubisco in *Arabidopsis* requires a capacity for redox regulation of the larger rubisco activase isoform. *Proc. Natl. Acad. Sci. 99*, 3330–3334. doi: 10.1073/pnas.042529999

(224) Zhang, P., Battchikova, N., Jansen, T., Appel, J., Ogawa, T., and Aro, E.-M. (2004) Expression and functional roles of the two distinct NDH-1 complexes and the carbon acquisition complex NdhD3/NdhF3/CupA/Sll1735 in *Synechocystis* sp PCC 6803. *Plant Cell 16*, 3326–3340. doi: 10.1105/tpc.104.026526

(225) Zimorski, V., Ku, C., Martin, W. F., and Gould, S. B. (2014) Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol. 22*, 38–48. doi: 10.1016/j.mib.2014.09.008

(226) Zmasek, C. M., and Godzik, A. (2013) Evolution of the animal apoptosis network. *Cold Spring Harb. Perspect. Biol. 5*. doi: 10.1101/cshperspect.a008649