Understanding Spatial and Temporal Organization of Cyanobacteria for Synthetic Biology Applications

Understanding Spatial and Temporal Organization of Cyanobacteria for Synthetic Biology Applications

A DISSERTATION PRESENTED
BY
ANNA HANG CHEN
TO
THE COMMITTEE ON HIGHER DEGREES IN SYSTEMS BIOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
SYSTEMS BIOLOGY

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
APRIL 2015

Understanding Spatial and Temporal Organization of Cyanobacteria for Synthetic Biology Applications

Abstract

The goal of synthetic biology is to engineer biological systems in order to solve industrial and medical challenges, as well as to learn about these systems by building. Cyanobacteria, a chassis for such engineering, are major players in the global carbon cycle and their ability to fix carbon has been harnessed to produce various chemicals, including biofuels. In addition, cyanobacteria possess remarkable spatial and temporal organization in the cell. In this dissertation, I monitor, break down, and rebuild the molecular components necessary for this spatial and temporal coordination of cyanobacterial growth. These studies give us a better understanding of basic cyanobacterial biology and enable the further development of cyanobacteria for synthetic biology applications. In chapter 1, I describe the arrangement of the cyanobacterial chromosomes over time, showing the mechanisms regulating chromosome duplication and segregation. The polyploid nature of cyanobacteria make this study relevant to their efficient genome engineering. In chapter 2, I elucidate the assembly of the primary carbon fixation machinery, carboxysomes. I show that the internal cargo of carboxysomes, RuBisCO, seeds assembly, followed by the recruitment of shell proteins, which form a solvent-protected microenvironment. Finally, in chapter 3, I engineer a synthetic circadian clock from cyanobacterial components in a heterologous organism, *E. coli*. I demonstrate the clock's modularity and pave the way for its use in medical and

industrial applications. Taken together, this work furthers our understanding of cyanobacterial physiology and forms a foundation for their efficient engineering to increase their carbon fixation capabilities. Furthermore, the fundamental spatial and temporal organization strategies elucidated here can serve as inspiration for the engineering of heterologous systems that can serve similar purposes in different contexts.

Contents

Listing of figures

Copyright Information

Acknowledgments

First and foremost, I would like to thank my advisor, Pam Silver, for her support throughout this journey. She welcomed me into a lab full of smart talented people (twice!), provided me with the resources and guidance to pursue a vast array of projects, and encouraged me to think big and achieve more.

I owe much of what I've accomplished to Dave Savage, who taught me how to look at scientific problems with a rigorous but also curious mind, showed me the ropes on a variety of experimental techniques, and helped me with my scientific writing. He mentored me starting from when I was a clueless undergraduate and has always been there to offer advice and support whenever I needed it. I would also like to thank my two baymates, Danny Ducat and Roger Chang, who have taught me so much about science and life in general. Thanks to Jessica Polka for being an awesome collaborator on the carboxysome project and to David Lubkowicz and Vivian Yeong, who worked with me on the circadian rhythm project. I would not have been able to take those projects as far as they went without their help. Thanks to Joe Torella for interesting late night lab conversations and for guidance on just about everything. To all my friends in the Silver Lab: Avi Robinson-Mosher, Mara

Inniss, Tyler Ford, Steph Hays, Matt Mattozzi, Drew MacKeller, Marika Ziesack, Bruno Afonso, and everyone else, thanks for keeping the lab a lively, fun and enjoyable space.

I would also like to acknowledge Padraig Deighan, my collaborator on the two-hybrid experiments, as well as Nate Lord, my collaborator on the mother machine experiments.

I would like to thank my dissertation advisory committee: Peng Yin, Jim Collins, Ann Hochschild, and Roy Kishony. They provided input and feedback on my projects over the years. Thanks also to my defense committee: Peng Yin, Ann Hochschild, Ann Pearson, and Dan Needleman.

I could not have made it this far without my friends, who keep me sane when science isn't going well and cheer me on when it is. Thanks to all my high school and college friends, for keeping me nerdy. To all my wonderful friends on the ballroom team, especially Lihua Bai, Mark Chen, Eddie Kay, Armen Mkrtchyan, and Aaron Beal, thanks for dancing with me and for being patient when I was stuck in lab. Thanks to Jenn Brophy and Susan Chen for our awesome JAGS meetings.

Thanks to Gabriel Wu for pushing me to be better, for caring about me in every way, and for all of our awesome adventures. I am truly lucky to have met you.

Finally, I would like to thank my family. Thanks to my parents, Amy Liu and James Chen, for their love and support. They have always given me the confidence to be the best I can be and provided me with help and guidance whenever I needed it. Thanks also to my siblings, Amanda Chen and Alex Chen, for always being able to make me laugh and for being great friends.

0 Introduction

0.1 Cyanobacteria

Cyanobacteria are photosynthetic bacteria that play a key role in the global carbon cycle[43]. *Synechococcus elongatus PCC 7942* are freshwater cyanobacteria and obligate photoautotrophs, previously known as *Anacystis nidulans R2*[177]. They are rod-shaped, gram-negative bacteria, about 2-4μm in length and about 0.5 μm wide[148]. *S. elongatus* is naturally competent and has a fully sequenced 2.7Mb circular genome[69]. It is a model organism for the study of carbon fixation, photosynthesis, and bacterial circadian rhythms.

Cyanobacteria have been proposed to be useful in a wide variety of applications,[149] from production of biofuels to detoxification of waste water. A better understanding of their fundamental systems will facilitate their design and engineering. The advantage of using

cyanobacteria for carbon sequestration into industrially relevant compounds is that they can grow on marginal agricultural land and require relatively cheap inputs[32]. Wild-type cyanobacteria already produce many natural products, including sugars and isoprene[32,104]. Furthermore, cyanobacteria have been engineered to produce a variety of biofuels, including alcohols, alkanes and hydrogen[104,34]. Other commodity chemicals, which are used in pharmaceuticals, flavors, and fragrances, have also been produced in cyanobacteria[176]. In addition, strategies to secrete these products have been developed[31] in order to increase yields and decrease costs of production. A better understanding of cyanobacterial biology as well as the ability to engineer cyanobacteria to increase biomass yields from carbon fixation are necessary to enhance our ability to harness photosynthesis and produce useful chemicals out of inorganic carbon in a carbon neutral manner.

Despite their potential, cyanobacteria are not commonly used in industrial applications. The major challenge is one of economic feasibility—production in cyanobacteria is not economically viable unless the compound is of high value. Currently, cyanobacterial biomass production costs are estimated to be €4.95/kg, which is equivalent to €162.79/GJ of energy[124]. A sensitivity analysis performed by Norsker *et al.* showed that further optimization of plant location, media conditions, and bioreactor design could bring the cost down to €0.68/kg, which would be €25/GJ[124]. While this cost is on par with the cost of delivered electricity, it is still higher than that required to compete with fossil fuels (€5/GJ). Thus, in order to make cyanobacterial production of biofuels and other commodity chemicals viable, a further increase in photosynthetic conversion efficiency must be accomplished. Currently, the solar energy to biomass conversion rate is <1% whereas the theoretical maximum energy conversion is approximately 4.5%[4]. An increase in photosynthetic efficiency

to achieve a level close to the theoretical maximum would accomplish much of the needed cost reduction to make industrial use of cyanobacteria viable. To that end, organism and metabolic engineers would traditionally use standard genetic engineering tools to optimize pathways and balance flux. While this is possible for cyanobacteria, there remain unique challenges that, if solved, would greatly facilitate the engineering process.

One obstacle to engineering cyanobacteria is their polyploid genome[67]. While cyanobacteria are naturally competent and therefore readily uptake DNA[42], individual cells contain between 1-10 copies of their chromosome[17]. Exogenous gene expression by recombination into the genome is routinely done, but the presence of additional unmodified genome(s) in cyanobacteria require higher costs to maintain expression due to the need for selective pressure to maintain the modified chromosomes. Creating a complete knockout strain is a challenge as well, since it also requires proper selective pressure to ensure elimination of a gene from all copies of the genome. Once a better understanding of the polyploid genome and its duplication and segregation properties is achieved, chromosomal manipulation of cyanobacteria for industrial use would be easier than traditionally less competent organisms.

Another major factor limiting production in cyanbacteria is the conversion rate of carbon dioxide into biomass and other more useful chemicals. Evolutionarily, cyanbacteria and other photosynthetic organisms evolved carbon fixation while carbon dioxide was plentiful and oxygen was scarce[3]. As a result, the primary carbon fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), is inefficient and slow, with a turnover rate of about 3/sec[175]. In addition, RuBisCO also catalyzes a waste oxygenase reaction, unable to distinguish carbon dioxide from oxygen[175]. This reaction produces a waste prod-

uct that the cell must then recycle back into usable carbon. In the process, a previously fixed molecule of carbon dioxide is released[39]. While plants have evolved to produce a large amount of RuBisCO, totaling over 50% of its biomass[40], cyanobacteria have evolved a carbon concentrating mechanism utilizing a bacterial protein microcompartment[84], known as the carboxysome. Understanding the formation and development of the carboxysome will allow production optimization strategies that utilize intracellular organization to improve yields.

Cyanobacteria are robust (highly tolerant to environmental insults), autotrophic, genetically tractable organisms that show great promise for useful applications in industrial processes. Several challenges must still be overcome before they are economically viable, but understanding their spatial organization, from genomes to carboxysomes, will provide a roadmap to achieving industrial relevance.

0.2 Intracellular Organization*

0.2.1 The need for spatial organization

Cells, including but not limited to cyanobacteria, face many challenges that benefit from spatial organization (Figure 0.1a). First, some enzymes, such as RuBisCO[38], suffer from slow turnover, which results in flux imbalances or bottlenecks in metabolic pathways. Reliance on such enzymes may require establishing high local substrate concentrations to achieve reaction rates high enough to support adequate pathway flux[81]. Second, diffusion of volatile intermediates through the cell membrane results in their loss from the cell[108].

*Portions of this chapter are reproduced from the following work: Chen, A. H., Silver, P. A. (2012). Designing biological compartmentalization. *Trends in Cell Biology*, 22(12), 662–70. Reprinted with permission from Elsevier.

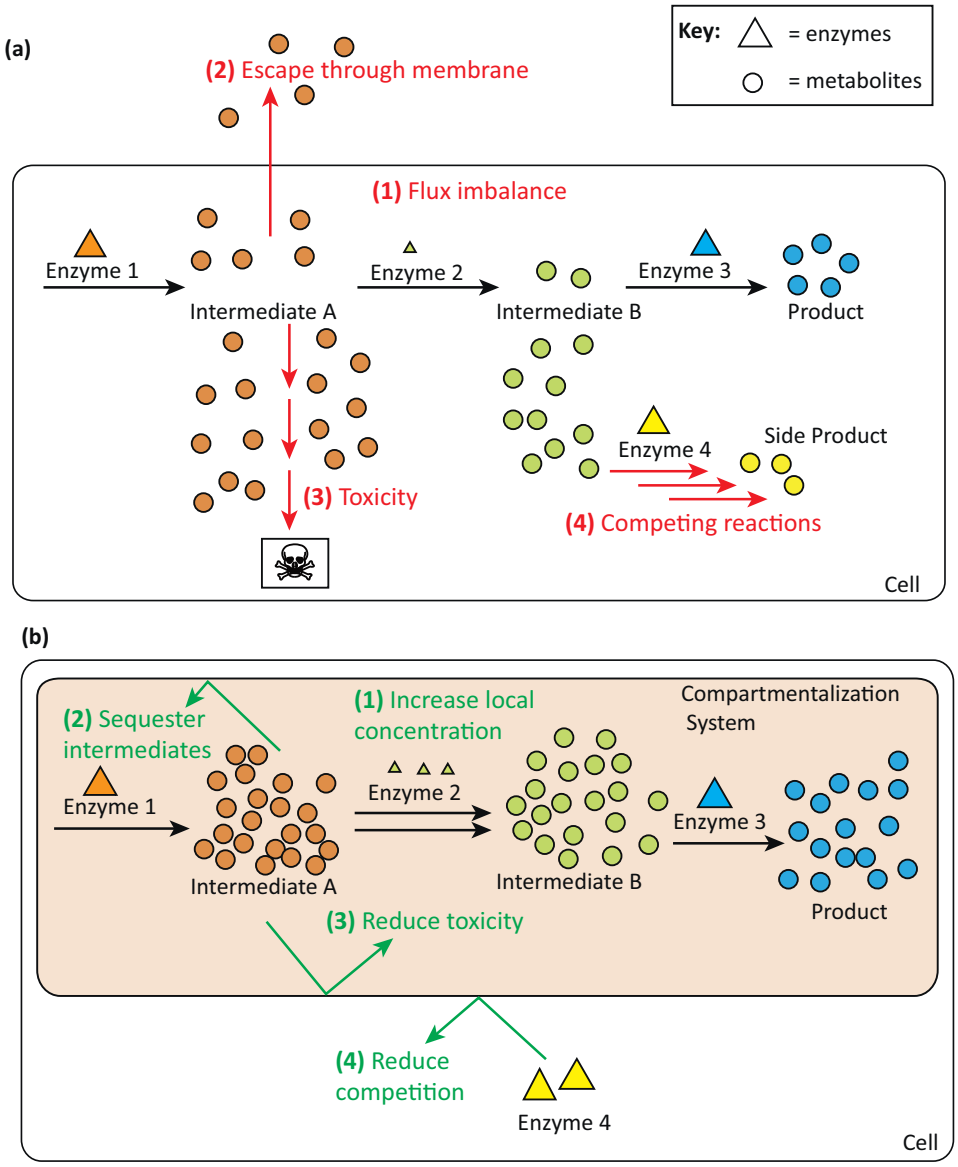

Figure 0.1: Benefits of Spatial Organization. (a) Nature faces many challenges when conducting the chemical reactions of the cell. (1) Differing enzyme kinetics may result in flux imbalances. (2) Intermediates may be lost through the cell membrane. (3) Toxic intermediates can result in growth inhibition. (4) Competing reactions can divert flux through undesired pathways. (b) Compartmentalization systems specifically solve challenges 1–4, respectively, by: (1) creating areas of local concentrations to favor reaction kinetics; (2) sequestering intermediates; (3) reducing toxicity; and (4) reducing competition.

Third, biosynthetic pathways can generate toxic intermediates that inhibit growth, such as hydrogen sulfide accumulated during bacterial sulfur metabolism[99]. Finally, metabolites can participate in multiple competing reactions, reducing their availability for any single pathway. An example of this is malonyl-CoA, an intermediate that is consumed in fatty acid and phospholipid production but is also used in the biosynthesis of polyketides and flavonoids[182].

To deal with these challenges, nature has evolved compartmentalization strategies (Figure 0.1b), such as large enzyme complexes[71,24,131] and organelles[110,83], to spatially organize metabolism. In eukaryotes, compartmentalization in the form of membrane-bound organelles is common. The peroxisome, for example, encapsulates reactions that generate or consume hydrogen peroxide, a toxic intermediate from the breakdown of organic substrates in oxidative reactions[53].

Until recently, prokaryotes were generally thought to lack internal organization[103]. With the exception of a few rare examples, including pirellulosomes[98] and magnetosomes[116], prokaryotes generally do not contain membrane-bound internal compartments. However, researchers have recently discovered different types of bacterial microcompartments that partition the internal space of the bacterial cell for specialized functions[184,83]. These proteinaceous shells have been found in approximately 400 microbial genomes[83]. For example, two different bacterial microcompartments protect cells from toxic aldehyde intermediates. The ethanolamine utilization (Eut) microcompartment sequesters acetaldehyde, a volatile and toxic intermediate of the ethanolamine utilization pathway[130]. Likewise, the 1,2-propanediol utilization (Pdu) microcompartment encapsulates propionaldehyde, minimizing its toxicity[142]. Another example of a bacterial microcompartment is the car-

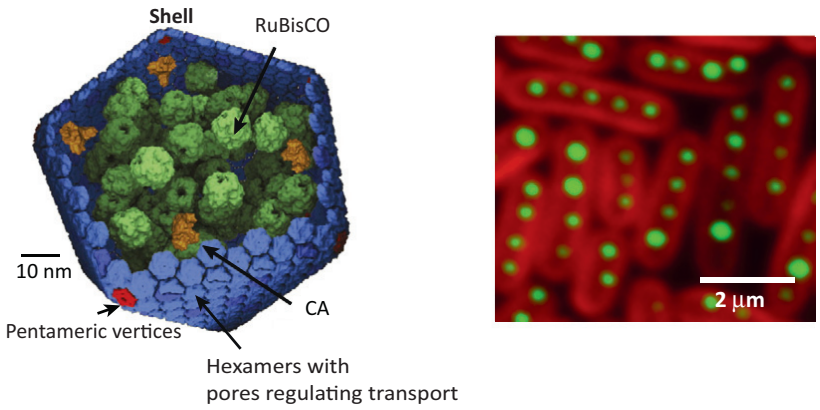

Figure 0.2: Bacterial Microcompartments are protein structures that encapsulate reactions. Carboxysomes contain ribulose 1,5-bisphosphate carboxylase oxygenase (RuBisCO) and carbonic anhydrase. The shell comprises pentamers and hexamers, some of which contain pores that are thought to regulate substrate transport (left). Carboxysomes spatially align in the cell (right). Heterologous expression of microcompartments and further engineering of targeting and transport has the potential for making microfactories that may be capable of encapsulating various reactions. (Left: reprinted [9] with permission from the authors. Right: image courtesy of David F. Savage, Berkeley, CA, USA).

boxysome. In all cyanobacteria and some chemoautotrophs[184], carboxysomes encapsulate RuBisCO and carbonic anhydrase, enzymes involved in the rate-limiting step of the Calvin cycle[162,13] (Figure 0.2). These microcompartments, rather than serving to protect the cell, are instead the primary 'carbon-concentrating mechanism' in these bacteria, facilitating a reaction that would otherwise be difficult to sustain. Carboxysomes are proposed to help overcome the slow turnover rate of RuBisCO by providing a high local concentration of carbon dioxide to the enzyme[163,3].

0.2.2 Synthetic compartmentalization

Synthetic biologists have drawn inspiration from nature to design synthetic compartmentalization systems. Many of the biosynthetic pathways that we want to engineer present the same challenges of toxic intermediates, competing reactions, and flux imbalances found in nature[24,81]. Therefore, biologically inspired engineering of synthetic compartmentalization can solve some of these challenges. Synthetic compartmentalization strategies include protein scaffolds[33], nucleic acid scaffolds[27,48,25], and lipid compartmentalization[6,45], all of which have been shown to increase production for metabolic engineering purposes.

In addition, a few microcompartments have been heterologously expressed and could be engineered to encapsulate foreign enzymes. For example, a lumazine synthase capsid from *Aquifex aeolicus* was heterologously expressed in *E. coli*[180]. These capsids were engineered to encapsulate a toxic enzyme that was produced cytoplasmically, HIV protease. The sequestering of HIV protease resulted in a growth advantage. Thus, the size, shape and assembly properties of the capsid were modifiable by directed evolution or by rational design and engineering[18]. In contrast, most other synthetic microcompartments are not yet customizable. The Eut and Pdu microcompartments from Salmonella have been expressed in *E. coli* in their original form[20,127]. Also, carboxysomes from cyanobacteria (Figure 0.2) were heterologously expressed in *E. coli* and the encapsulated RuBisCO was functional in an *in vitro* assay[9].

While some progress has been made in building synthetic microcompartments, many challenges remain before they can be utilized in chemical production and other applications. First, bacterial microcompartments are relatively closed to the surrounding environment, thus the microenvironment created inside must be well characterized before the

microcompartment can be used in a production setting. Secondly, microcompartments contain protein-based pores that likely regulate transport of substrate into and out of the proteinaceous shell (Figure 0.2)[90]. These must be engineered to carry the substrate of interest. Lastly, and most importantly, targeting of foreign enzymes to microcompartments is an area of ongoing research. Localization sequences have been found for the Pdu and Eut microcompartments[20,44], but the targeting novel enzymes to the carboxysome remains poorly understood[9]. While associations between carboxysome shell proteins and other protein components have been shown[86], the mechanism of assembly and the properties necessary for this self-assembling microcompartment remains unclear.

Overall, protein assemblies have been used effectively to organize metabolic reactions and increase product yields. Thus far, however, such increases have been low. Therefore, the present challenge is to further increase yields to levels high enough for industrial applications. A better understanding of the mechanisms underlying cellular spatial organization strategies will aid in meeting this challenge, because we may be able to engineer the geometry, selectivity and assembly properties to more robustly increase reaction rates and decrease toxicity and leakage. In addition, scaffolds have been applied to only a few select pathways as proof of principle. Generality and scalability of scaffolds will require a better understanding of localization to synthetic scaffolds. Herein, we expand our knowledge of compartmentalization strategies of the cell towards the goal of making it widely applicable to increasing yields in metabolic engineering efforts. In addition, implementing and studying synthetic systems will help us understand the native biological spatial organization of the cell.

0.3 Temporal Organization

"As physical beings, individual cells and whole organisms are bound by the four dimensions occupied by space and time....A full understanding comes from knowledge of the interplay between space and time."[1]

In addition to spatial organization, timing is important for many biological processes, including development[87] and metabolism[82,136]. A variety of natural clock systems have evolved in various organisms to keep track of biological time. For example, heartbeats occur at regular intervals. Annual rhythms occur in mammals that hibernate, migrate and breed[89] . On the cellular level, cell cycle timing is regulated by a series of checkpoints[117]. In mammalian cells, the Hayflick limit determines the lifespan of a cell before it reaches senescence[150]. Metabolic rhythms oscillate on a short time scale (hours to minutes) and regulate fundamental processes in the cell. For example, yeast *Saccharomyces cerevisiae* have 40 minute respiratory cycles, as well as 4 to 5 hour ultradian cycles[118]. The most commonly studied, and one of the most widely prevalent biological timing mechanisms is the circadian rhythm[147], which allows organisms to time various events in coordination with the 24 hour day-night cycles of the earth.

Circadian rhythms are ubiquitous across all kingdoms of life[37] and regulate many activities. These range from fundamental molecular processes such as gene expression[158,91], to whole organism phenotypes such as flowering[126]. In humans, abnormalities in circadian rhythms are present in a wide array of disorders, including insomnia and even depression[55]. Recently, it was even discovered that disruption of the circadian rhythm of the human gut microbiome can result in obesity and metabolic imbalances in the host[165]. Furthermore this

phenotype was directly transferrable via the microbiome, suggesting it is independent of other host factors[165]. In all, a proper circadian rhythm is a crucial element of human health and has potentially severe consequences when disrupted[181].

Circadian rhythms have three fundamental properties[140]. The first is that they are free-running, that is, they must be able to maintain a 24 hour period independent of external stimuli (such as light-dark cycles). The second is that they must be entrainable. If a rhythm is out of phase with the external environment, light or other inputs, known as "zeitgebers", should be able to shift the rhythm back in synchrony with its surroundings. The third property is temperature compensation. Temperature compensation prevents shifts in temperatures, such as those present in seasonal variations, from dramatically affecting the circadian rhythms.

The study of biological timing, chronobiology, has uncovered the mechanism of action of many biological timing devices. In humans, the suprachiasmatic nuclei (SCN), a group of cells located in the hypothalamus is the master circadian clock[26]. The SCN regulates many peripheral clocks in various organ systems, including the lungs, liver, spleen, and skin[28]. The molecular mechanism driving the human circadian clock involves complex transcriptional-translational feedback loops among a series of genes[50].

In contrast to the complex circadian clock found in humans, cyanobacteria have a relatively simple clock. At the core, the clock consists of just three proteins, KaiA, KaiB, and KaiC[74]. It was found that these three proteins isolated *in vitro*, along with added ATP, can result in post-translational circadian oscillations in the phosphorylation state of KaiC[119]. This simple clock controls the expression of the majority of cyanobacterial genes[100], resulting in day-night changes in photosynthesis, metabolism and other processes. However,

neither the inputs into nor the outputs from this clock are well understood. Given the far reaching effects of the circadian clock, a better understanding of its mechanism of action could allow us to harness its abilities and engineer a higher producing or better growing strain of cyanobacteria.

0.3.1 The need for synthetic timing

Unlike strategies involving spatial organization, which synthetic biologists have attempted to mimick in various synthetic devices, there are very few examples of natural timing strategies that have been built in synthetic systems. However, there is a need for synthetic timing devices. This need is two fold. First of all, due to the complex nature of biological timing, it is difficult to study in its native environment. Synthetically constructing biological clocks will allow us to study their core compositions and mechanisms of action in isolation. Examples of this type of "learning by building" can be found in the heterologous expression of various biosynthesis pathways[179]—when put in a heterologous organism, the inherent feedback systems and the underlying mechanisms of each step in a pathway can be studied by observing their level of functionality in this new, orthogonal context. Secondly, a synthetic timing device would be useful for a wide array of applications. Chronotherapy, the strategic timing of treatment with diurnal rhythms, has been used to treat affective disorders as well as to improve the treatment efficacy of chemotherapy[157,95]. With an increasing understanding of chronotherapeutics as well as the advent of novel complex therapeutics, the need for timing devices to complement drug delivery and other therapies is increasing. Already, synthetic biology research is attempting to use logic gates and other control devices to supplement new therapeutics, such as CAR T-cell therapies[15]. An additional timing

dimension could help control drug delivery timing, alter dosage duration or sense other temporal metrics in the body.

0.3.2 Synthetic timing devices

Synthetic biologists have built many synthetic circuits, such as toggle switches, multi-input RNA regulatory systems[19], and others reviewed by Brophy *et al*[10]. Most of the extant devices are designed as biological computation machines. A subset, however, focuses on mimicking natural timing processes. On a fundamental level, transcriptional elements have been used to design time delays. Biological rhythms add another layer of complexity, requiring robustness and entrainability. There have been two major attempts at mimicking biological cycles using synthetic devices. One is the repressilator, a network of three transcriptional repressor systems that together induce periodic synthesis of green fluorescent protein (GFP) with periods of hours[41]. The second timing system is a tunable mammalian synthetic oscillator[166]. The period of this oscillator is short (minutes) and not biologically relevant to the cell or circadian cycle. Both the repressilator and this mammalian oscillator keep time in a biological system, but both were rationally designed using mathematical models and do not utilize components from any natural biological clock.

Therefore, while synthetic spatial organization that mimicks biology has been implemented, synthetic timing devices are still in their early stages of implementation. Elowitz states in his work: "It would be interesting to see whether one could build an artificial analogue of the circadian clock"[41], indicating the need for a synthetic oscillator with longer periods and a biological analogue. Thus far, the devices that have been built have been used as tools for understanding biological timing, but due to their artificial nature, are ul-

timately poor models for physiology studies. Additional devices, with biologically inspired designs and biologically relevant periodicities, should be built in order to better understand chronobiology as well as for use in medical and industrial interventions.

0.4 Summary

In order to better harness the many applications that cyanobacteria can have, it is necessary to engineer them. To do that, we must first understand cyanobacterial biology and the spatial and temporal organization in cyanobacterial cells. In chapter 1, I study cyanobacterial genomes and visualize their duplication and segregation. In chapter 2, I describe a study of the cyanobacterial carboxysome, the crucial carbon fixing organelle. I study the assembly of this microcompartment as well as the microenvironment inside the shell. In Chapter 3, I build a heterologous circadian oscillator and demonstrate its functionality. Together, the work presented in this dissertation represents significant progress towards the understanding of cyanobacteria physiology and paves the way for better engineering of these organisms.

1

Spatial and Temporal Organization of Chromosome Duplication and Segregation in Synechococcus elongatus PCC 7942

1.1 ABSTRACT*

The spatial and temporal control of chromosome duplication and segregation is crucial for proper cell division. While this process is well studied in eukaryotic and some prokaryotic organisms, relatively little is known about it in prokaryotic polyploids such as *Synechococcus elongatus PCC 7942*, which is known to possess one to eight copies of its single chromosome. Using a fluorescent repressor-operator system, *S. elongatus* chromosomes and

*Portions of this chapter are reproduced from the following work: Chen, A. H., Afonso, B., Silver, P. A., & Savage, D. F. (2012). Spatial and temporal organization of chromosome duplication and segregation in the cyanobacterium Synechococcus elongatus PCC 7942. *PLOS ONE*, 7(10), e47837. doi:10.1371/journal.pone.0047837. Reprinted with permission under the Creative Commons Attribution (CC BY) license.

chromosome replication forks were tagged and visualized. We found that chromosomal duplication is asynchronous and that the total number of chromosomes is correlated with cell length. Thus, replication is independent of cell cycle and coupled to cell growth. Replication events occur in a spatially random fashion. However, once assembled, replisomes move in a constrained manner. On the other hand, we found that segregation displays a striking spatial organization in some cells. Chromosomes transiently align along the major axis of the cell and timing of alignment was correlated to cell division. This mechanism likely contributes to the non-random segregation of chromosome copies to daughter cells, resulting in most daughter cells inheriting half of the chromosomes of the mother cell.

1.2 Introduction

Genomic DNA replication and segregation are fundamental processes crucial to survival for all organisms. This process has been well studied in many bacterial species, including *Escherichia coli*[92], *Bacillus subtilis*[7,168,60], and *Caulobacter crescentus*[64,170]. Most of these organisms possess a single copy of one, two or three different chromosomes (Figure 1.1A, I-III). In contrast, the cyanobacterium *Synechococcus elongatus PCC 7942* has multiple copies of its single chromosome—estimates suggest between three to six copies[61,8] (Figure 1.1A, IV). To date, little is known about the dynamics of replication and segregation in prokaryotes with multiple copies of a single chromosome.

Most studies of replication have been conducted in monoploid bacterial species. In many of these organisms, replication timing and synchrony is strictly regulated[114]. In *E. coli*, for example, all origins fire synchronously at a fixed cell size per origin (initiation mass) that is independent of growth rate[64,101]. Synchrony is tightly coupled to cell division cycles

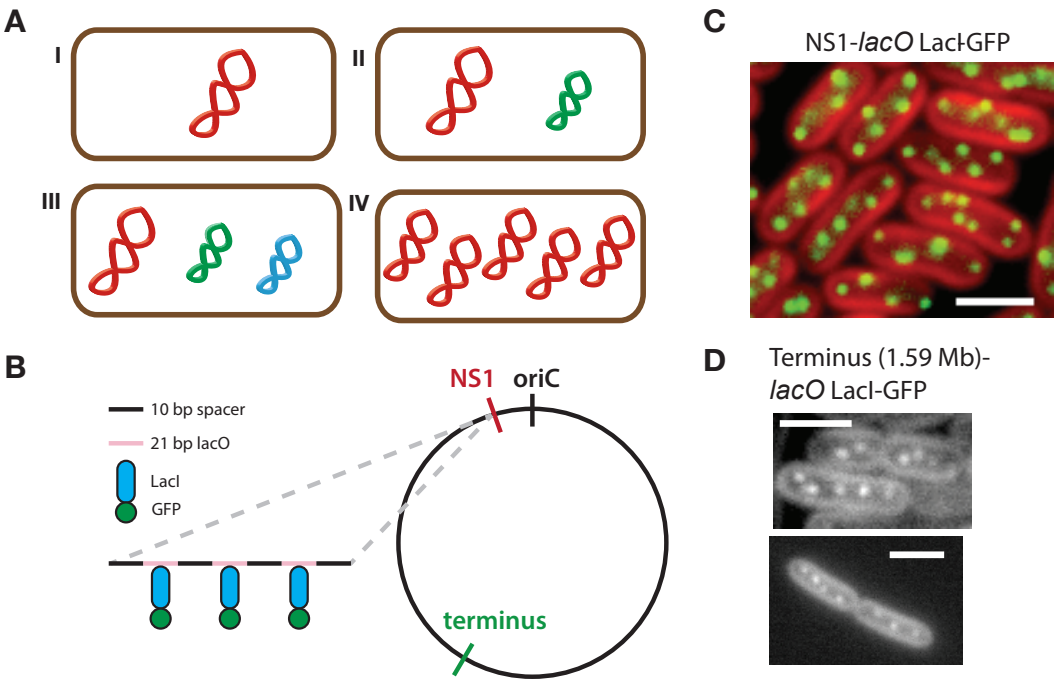

Figure 1.1: Chromosomes in the polyploid bacterium *S. elongatus* can be visualized using a fluorescent repressor-operator system. (A) Bacteria contain different genomic arrangements. Here, each color represents a different chromosome. They can possess a single copy of one chromosome (I), or have multipartite genomes (II-III) with one large chromosome (red) and one or more smaller chromosomes (green and blue). Some species of bacteria, such as cyanobacteria *Synechococcus elongatus PCC 7942*, are polyploid. That is, they have multiple complete copies of one chromosome (IV). (B) Chromosomes can be tagged and observed *in vivo* using a fluorescent repressor-operator system. *lacO* arrays were integrated either near the origin of replication (NS1) or the predicted terminus in the *S. elongatus* chromosome. 10 bp spacers with random sequences were inserted between the operator sites to avoid recombination (black). The protein fusion LacI-GFP (blue and green) bound to multiple repeats of its cognate *lacO* operator site (pink). (C) The fluorescent repressor-operator system from (B) was transformed into *S. elongatus*. The origins of replication of each chromosome appear as foci (green) in cells (red) when imaged using wide-field fluorescence microscopy. Origins of replication are seen throughout the cell. (D) Cells with *lacO* arrays integrated near the putative terminus region at 1.59 Mb in the genome were visualized. Foci appear throughout the cell, similar to (C).

and ensures that daughter cells receive the correct number of chromosomes. However, regulating timing of replication may not be as important for proper cell division in polyploid organisms.

In addition to timing of replication, the spatial localization of replication is important for proper cell division. In *E. coli*, newly synthesized chromosomes appear in the cell center or at the quartile points along the long axis of the cell[138]. Replication forks appear at the origin of replication, separate into two sister replisomes that migrate to opposite cell halves as replication proceeds and returns to mid-cell as replication ends[138]. In *C. crescentus*, replisomes move towards the middle from the cell poles[78].

While some bacteria have eukaryotic mitosis-like segregation mechanisms[47], it has been suggested that organisms with multiple chromosomes do not require an active segregation mechanism, since given a large number of chromosomes, it is likely that each daughter cell will receive at least one copy[60]. This is analogous to high-copy plasmid systems, which typically lack an active segregation mechanism[35].

In order to better understand chromosome replication and segregation in the polyploid organism, *S. elongatus*, we tagged and visualized the chromosome and the replisome using a fluorescent repressor-operator system. Chromosome count and localization data was collected. We found that chromosome number correlates with cell length but that chromosome duplication timing is asynchronous. Thus, while duplication is associated with cell growth, it is not coupled to cell division. We also found that duplication occurs at random locations in the cell, but movement of each individual replisome remains confined after initiation. In addition, we probed the spatial organization of *S. elongatus* chromosome segregation and found that, contrary to previously suggested models, a surprising alignment

occurs during the process. Together, these results elucidate chromosomal replication and segregation dynamics in a polyploid prokaryote.

1.3 Materials and Methods

1.3.1 Bacterial Strains and Growth Conditions

The wild-type *Synechococcus elongatus PCC 7942* strain was acquired from the American Type Culture Collection (ATCC). *S. elongatus* cells were grown in solid BG11 medium following standard protocols with an illumination of 2000 lux at 30° C[2]. *S. elongatus* were transformed following standard protocols by incubating cells overnight in the dark with 100 ng of plasmid DNA and plating on selective media[21]. Antibiotics (kanamycin, spectinomycin, or chloramphenicol) were used at a concentration of 5 μg/ml. To prevent disruption of chromosome replication during growth, 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the media. Cells were then replica plated onto media with 50 μM IPTG[21] for visualization and further experiments.

1.3.2 Plasmid Construction

All cloning, unless otherwise stated, was done using a Biobrick-like strategy (SpeI as the upstream site and XbaI-HindIII-NotI as the downstream sites)[133]. 21 bp *lacO* operator sites were assembled with random ten bp spacers. *lacO* arrays were obtained from pLAU443[93]. Two *lacO* arrays, with 120 *lacO* sites each, were then assembled with a kanamycin resistance marker inserted between them. Using NheI and SalI restriction enzymes, this series of *lacO* arrays was then cloned into the neutral site 1 vector pAM2314[105] or a vector containing homology regions to the terminus at 1.59 Mb in the *S. elongatus* chromosome. In the same vec-

tor *lacI*, fused to either the superfolder variant of green fluorescent protein (*gfp*) or yellow fluorescent protein (*yfp*) was inserted.

1.3.3 Image Acquisition and Analysis

Cells were plated onto BG11+50 μM IPTG[21] +2% agarose pads, which were transferred to a glass bottom dish (MatTek) for imaging. A Nikon TE-2000 microscope with a 100×1.4 numerical aperture objective equipped with an ORCA-ER CCD camera was used. Image acquisition utilized custom software, written using MATLAB (Mathworks), which interfaced with the microscope control package μManager[159]. Lighting necessary for cell growth during time lapse microscopy was controlled via a network AC power controller (IP Power 9258T), which also interfaced with MATLAB. Image analysis was performed using ImageJ[144], custom software written in MATLAB using the Image Processing Toolbox, and MicrobesTracker[155]. Cells were segmented using phase contrast images and cell size was calculated. Chromosomes were identified as foci in fluorescent images and their location and number calculated. Tracking and segmentation were verified manually and corrected as necessary.

1.3.4 Single-Stranded-Binding (SSB) Protein Visualization

SSB protein genes were cloned from *S. elongatus*, fused to *mOrange*, and cloned into the neutral site 3 vector using methods described above. The resulting plasmid was transformed into cyanobacteria either alone or with the *lacI-lacO* plasmid and visualized using methods described above.

1.3.5 Fluorescently Labeled Nucleotides Incorporation and Imaging

Cells were grown in the presence of 0.3% pluronic F-68 to OD750 = 0.4. Pluronic F-68 concentration was then elevated to 3% and fluorescently labeled nucleotides tetramethylrhodamine-5–2′-deoxy-uridine-5′-triphosphate(Roche) were added at a final concentration of 3 μM. After growth to late log phase, cells were washed in PBS and imaged as described above.

1.4 Results

1.4.1 Fluorescent Tagging of the Genome Using a Repressor-operator System Reveals the Spatial Localization of Origins and Termini of Chromosomes *in vivo*

Organization of chromosomes has been studied in cyanobacteria using DAPI and fixed-cell staining methods[70,146]. However, these methods only give a static and low-resolution image of chromosome localization. In order to visualize and quantify chromosome dynamics *in vivo*, we used a fluorescent repressor-operator system (Figure 1.1B), which consists of fluorescently-tagged DNA-binding proteins (repressors) that bind to their cognate recognition sequences (operators). Operator arrays recruit many repressors, which appear as foci when imaged using fluorescence microscopy[58,97]. In our case, we used the LacI repressor fused to either yellow fluorescent protein (YFP) or the superfolder variant of the green fluorescent protein (GFP) as our DNA binding protein[128]. Simultaneously, an array of 240 *lacO* operator sites[92] was inserted using homologous recombination at various positions in the chromosome (Figure 1.1B).

The precise location of the origin of replication (*oriC*) and the terminus region (*ter*) was

predicted using the program Ori-Finder on the *S. elongatus* chromosome[52] and was recently confirmed with experimental data[172]. This analysis revealed that the origin of replication is located at the region defined as the start in the current chromosome sequence from NCBI. It is the intergenic region between dnaN and *ccbZp* and contains 11 *dnaA* boxes (consensus sequence TTTTCCACA)[52]. Interestingly, the *dnaA* gene, which is usually found near *oriC* in other species, was found elsewhere in the cyanobacterial genome (at 1.1 Mb). The terminus region was not clearly defined via either GC skew or base disparity[52]. Due to the highly recombinant nature of the *S. elongatus* genome, the GC skew plot does not display a clear V-shape typical of organisms such as *E. coli*[14]. We reasoned that the terminus would be close to the region with the highest peak of GC disparity (Figure 1.2, pink line) so we inserted the *lacO* array at 1.59 Mb in the chromosome.

Visualization of tagged *oriC* showed multiple distinct foci throughout the cell (Figure 1.1C). This is unlike observations of *oriC* localization in *E. coli* and *B. subtilis*, where *oriC* were replicated and maintained at the poles of cells[173,59]. Tagged termini also displayed multiple foci throughout the cell (Figure 1.1D), whereas termini in *E. coli* are known to migrate from poles to mid-cell during cell division[120]. Also, the origin and terminus in *E. coli* are spatially separate, confined to distinct regions of the cell[92]. We did not find such spatial specificity in *S. elongatus* (Figure 1.1D, bottom). Together, these data confirm earlier studies that *S. elongatus* does indeed have multiple copies of its chromosome throughout the cell cycle[113]. They also reveal that the cellular localization of the *oriC* and terminus in *S. elongatus* is distinct from that in other bacteria. The strains with the array integrated near the *oriC* showed a better signal-to-noise-ratio, so this strain was used in subsequent experiments.

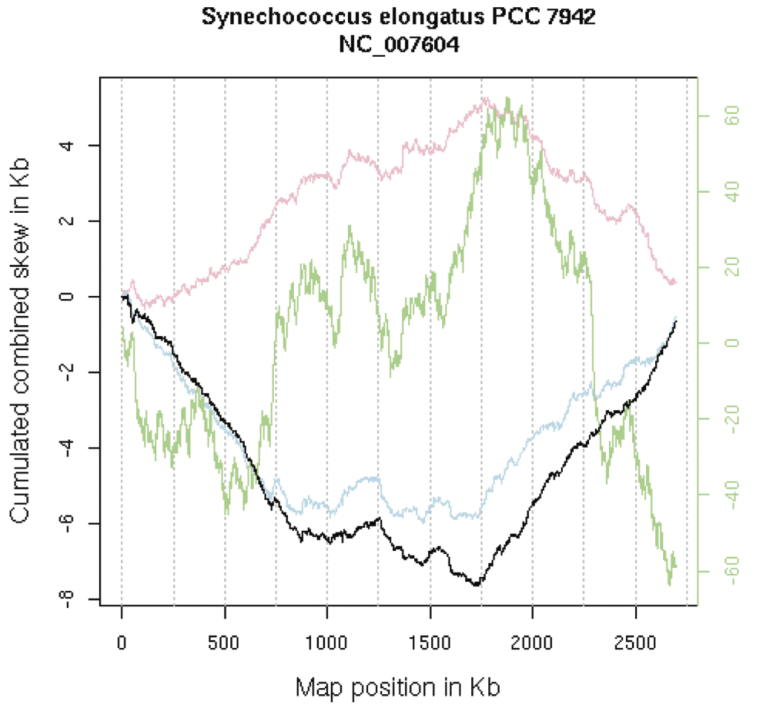

Figure 1.2: GC disparity mapped the terminus region of the *S. elongatus* chromosome. Due to its highly recombinant nature, the genome of *S. elongatus* gives rise to a GC-skew plot that does not display a clear V-shaped curve typical of organisms such as *E. coli* (black). For *S. elongatus*, the AT-skew (blue), the CDS skew (green), as well as the GC-skew (pink) is plotted. We reasoned that the terminus region would be present within the vicinity of highest peak of GC disparity (pink) so we looked for a region amenable for integration and inserted a 240 repeat *lacO*-array in a region located at 1.59 Mb in the chromosome.

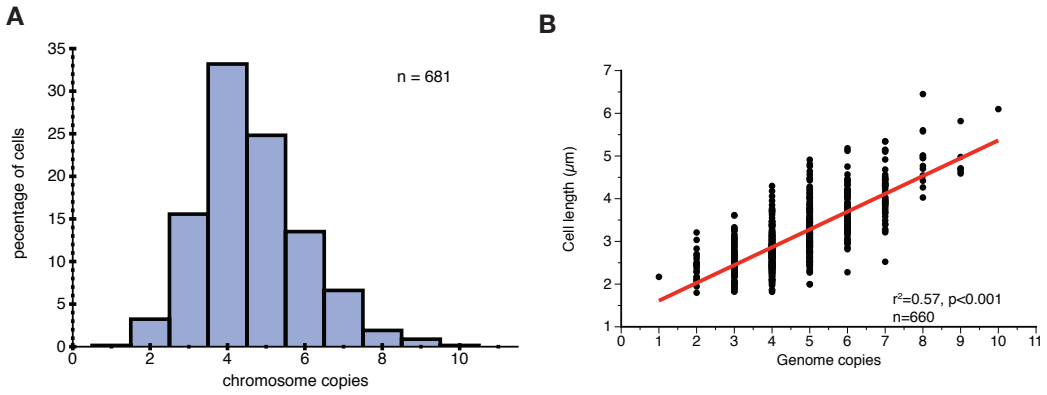

Figure 1.3: Chromosome duplication is correlated to cell length. (A) Distribution of chromosome number per cell is not significantly different from a log-normal distribution (n = 681, χ^2 goodness of fit test, h = 0, p = 0.2621). Most cells contain 4 chromosomes with values ranging from 1 to 10. (B) Chromosome copy number is correlated to cell length (n = 660, r = 0.7519, p<0.001), suggesting chromosome duplication is coupled to cell growth.

1.4.2 Chromosome Duplication is Correlated to Cell Length and not Coupled to Cell Division

By quantifying the number of fluorescent foci (tagged *oriC*), representing the number of chromosome copies in each cell, we found that the chromosome copy number distribution was not significantly different from a log-normal distribution (n = 681, χ2 goodness of fit test, h = 0, p = 0.2621) with a mean of 4.62 copies per cell and a median and mode of 4 copies per cell (Figure 1.3A). Cells harbored 1–10 chromosome copies. Interestingly, some cells contained an odd number of chromosomes and chromosome copy numbers other than 2n copies. This observation does not fit the model typically observed in single replicon prokaryotes, where replication occurs synchronously[114]. Instead, it is similar to observations of asynchrony in *E. coli* replication mutants[154]. The chromosome copy numbers hint at asynchronous DNA replication, supporting previous findings of constant DNA synthesis

rate over time; that is, DNA replication may not be coupled to cell division[113].

Indeed, we found that larger cells contained higher number of chromosome copies compared to smaller cells (Figure 1.3B), strongly supporting a model where chromosomes replicate at a constant rate during growth. By observing chromosome numbers in growing cells, we found a linear correlation between cell length and number of chromosomes (n = 660, r2 = 0.57, p<0.001) (Figure 1.3B, red line). These results suggest that chromosome duplication is correlated to cell growth.

1.4.3 Replication Timing is Asynchronous: Only One Replisome is found in Most Cells at Any Given Time

In order to better understand the timing and spatial localization of single replication events *in vivo*, we fluorescently tagged single-stranded-binding (SSB) proteins. SSB proteins play a fundamental role during chromosome replication, coating single-stranded DNA that is temporarily exposed; thereby, preventing it from degradation[63]. Approximately 30 SSB proteins localize to the replisome in *E. coli*, and this method has been used extensively to track replisomes in other organisms[5,137].

Cyanobacterial SSB protein was fused to mOrange and expressed in cells with the origin of replication tagged with GFP (as described above). We found that SSB foci (SSB-mOrange) were co-localized with Origin-GFP foci present in the cell (Figure 1.4A), indicating that tagging did not interfere with SSB function. SSB foci, therefore, correspond to active chromosome replication occurring in the cell. We found that at any given time, 85% of cells contain just one replisome, while 13.6% have two and only 1.3% have three (Figure 1.4B). Since most cells contain only one actively replicating chromosome (SSB-mOrange

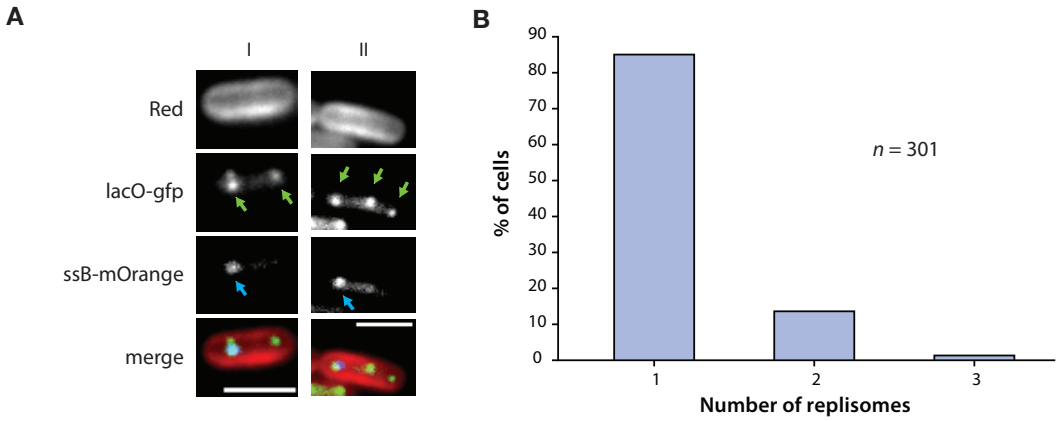

Figure 1.4: Chromosome duplication is asynchronous and is not coupled to cell division. (A) Single stranded binding (SSB) protein was tagged with mOrange. These were co-expressed in cells with LacI-GFP (*NSI-lacO*). Replisome localization appeared as foci and co-localized with tagged chromosomes (merge), indicating that tagging did not interfere with functioning of SSB. A cell with two chromosomes (left, green arrows) only contains one actively replicating chromosome (left, blue arrows). A cell with three chromosomes (right, green arrows) also only has one replisome (right, blue arrows). (B) Most cells contain one actively duplicating chromosome (85%), while the remaining contain two (13.6%) or three (1.3%) replisomes.

foci) but more than one chromosome (Origin-GFP foci), our data show that, in most cells, only one copy of the chromosome is being replicated at any given time. This, along with chromosome number data (Figure 1.3A) suggests that replication occurs asynchronously in *S. elongatus*.

1.4.4 Chromosome Duplication is Spatially Randomly Distributed

Since we found that cells containing multiple chromosomes typically replicate a single chromosome at a time (Figure 1.4), we hypothesized that new chromosomes may be synthesized at a particular location in the cell, either at the poles or mid-cell. We investigated the spatial localization of the replication event and of newly synthesized chromosomes to determine if there is a spatial preference for replication.

To accomplish this, we segmented cells and sub-segmented each cell along the major axis into 20 smaller regions. SSB foci location was distributed into these bins based on their distance from the pole of the cell along the major axis (Figure 1.5A). This distance was normalized to the total length of the cell. We found that the distribution of SSB localization was not significantly different from a uniform distribution (Kolmogorov-Smirnov test, h = 0, p = 0.4867, k = 0.0456), suggesting random localization of replisomes. That is, duplication is equally likely to begin at any point along the length of the cell in the inner quintiles (20%–80%). SSB at the poles of cells were not included in this analysis, since nucleoid volume results in reduced probability of the replisome appearing at the edge of the cell. In addition, SSB foci are less likely to be found at the poles due to decreased cell volume at the ends. From this data, we found a striking absence of spatial preference for beginning replication.

In order to confirm our finding that chromosome replication occurs in random locations throughout the cell, we also fluorescently labeled newly synthesized DNA. To do so, we permeabilized cells using pluronic F-68 to allow uptake of fluorescently labeled nucleotides[7]. After cells grew to late log-phase, incorporation of fluorescent nucleotides into the chromosome was seen as foci. The images were automatically segmented and binned as described earlier. There seemed to be no spatial preference in newly synthesized DNA (Figure 1.5B). Chromosome localization in the inner deciles (10%–90%) of the cell's major axis was not significantly different from a uniform distribution (KS test against uniform distribution. h = 0, p = 0.1027, k = 0.0579). We conclude that there is no spatial preference for chromosome duplication in the cell, confirming our previous results (Figure 1.5A) from replisome localization data. In contrast, *E. coli* chromosome duplication preferentially lo-

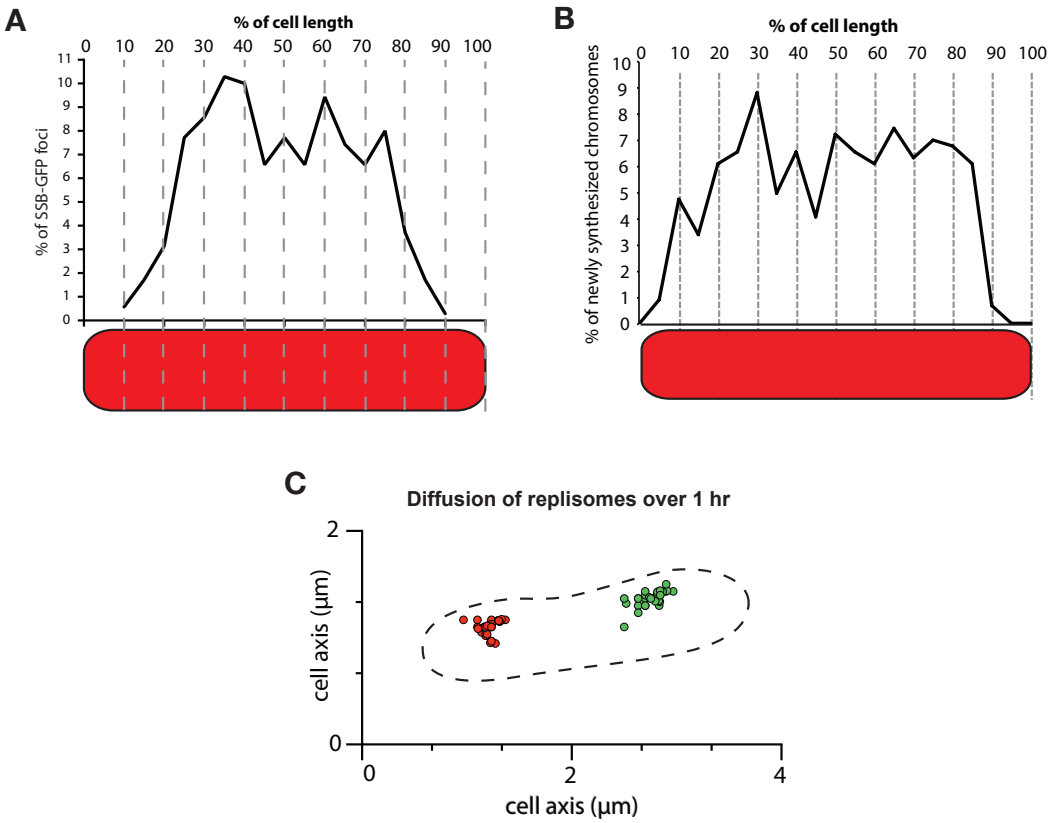

Figure 1.5: Replication occurs in a spatially random manner but replisomes undergo constrained movement. (A) Cells (n = 297) were sub-segmented along the major axis into 20 smaller regions and SSB foci localization was binned. The distribution of SSB localization was not found to be significantly different from a uniform distribution (KS test, h = 0, p = 0.4867, k = 0.0456). Cell poles were excluded in this analysis because of boundary effects due to reduced cell volume and the volume of the nucleoid decreasing the likelihood of SSB foci found at the edges of rod shaped cells. (B) To confirm the random localization of chromosome duplication, newly synthesized chromosomes were visualized using fluorescent nucleotide incorporation. Fluorescent foci also show a uniform distribution (KS test, h = 0, p = 0.1027, k = 0.0579), confirming the results shown in (A). (C) Two replisomes (red and green) were tracked over the time scale of a complete replication cycle with images taken every two minutes. Replisomes show restricted occupation of domains in the cytoplasm, remaining close to their initial position.

calizes to the center or the quartile points along the major axis (in cells with one and two replisomes respectively). In other organisms, chromosomes preferentially replicate at the poles or the middle of the cell[121,80,7]. Our findings on the randomly localized cyanobacterial chromosome duplication suggest a model that differs from existing models of chromosome duplication in other organisms.

1.4.5 Replication Fork Movement is Constrained

We concluded that replisomes appear at random locations throughout the cell, but how does their localization change over the course of a single duplication process? To investigate dynamic behavior of replisomes, time lapse imaging of SSB foci was used to track movement of replication forks. We followed replisomes over time, tracking individual foci every 5 seconds. We calculated a diffusion coefficient of $6.24\times10^{5}\pm4\times10^{5}$ $\mu m^2/s$ (n = 20) for replication forks, similar to the reported values for *E. coli* (10^4 $\mu m^2/s$)[137,138].

In order to determine whether movement is constrained over the time scale of a complete replication cycle, we tracked replisomes for one hour, acquiring images every two minutes (Figure 1.5C). We found that replisomes were confined to a region of the cytoplasm, remaining close to their initial position. The small amount of movement over the time scale of replication suggests that replisome motion is constrained to the local cytoplasmic region in which it assembled. This is distinct from *E. coli* or *C. crescentus* in which replication follows directed motion between mid-cell and the edges of the cell. Our data suggest that chromosomes are spooled through the replisomes rather than replisomes tracking along the chromosome. Together, these data give us an understanding of the spatial organization of chromosome duplication in *S. elongatus*: duplication starts in random locations through-

out the cell and as duplication proceeds, the replisomes remain stationary.

1.4.6 Chromosomes Transiently Align During the Cell Cycle

The lack of spatial preference of duplication events supports the current model that organisms with multiple chromosome copies do not have an active chromosome segregation system[60]. We surmised that our GFP-LacI-*lacO* method of visualizing chromosome dynamics *in vivo* would provide further insights into the dynamics of multiple chromosomes during segregation and allow us to follow up on previous works[146]. We found striking spatial organization, contrary to earlier models of random chromosome segregation in polyploid organisms.

Cells in a freely growing population displayed one of two phenotypes. Most of the cells displayed randomly localized chromosomes (85%, n = 289) (Figure 1.6A right). However, some cells displayed chromosomes aligned along the major axis of the cell (Figure 1.6A left). This surprising behavior prompted us to analyze the spatial arrangement of the chromosome copies over time. We tracked growing cells every hour for eight hours after which we could no longer distinguish signal from background (Figure 1.6B). We found that chromosomes transiently aligned during the cell cycle: collapsing towards the middle and aligning evenly spaced along the major axis of the cell (Figure 1.6B, 4 hrs, yellow arrows). Shortly after (1 hr), spatial arrangement was lost. This transient spatial arrangement took place either one or two times during the acquisition window of 8 hours. Occasionally, partial alignment was observed; that is, not all of the chromosomes aligned. This process was correlated to cell division, hinting that this process is mainly driven by the cell's commitment to division and may help maintain high fidelity chromosome segregation.

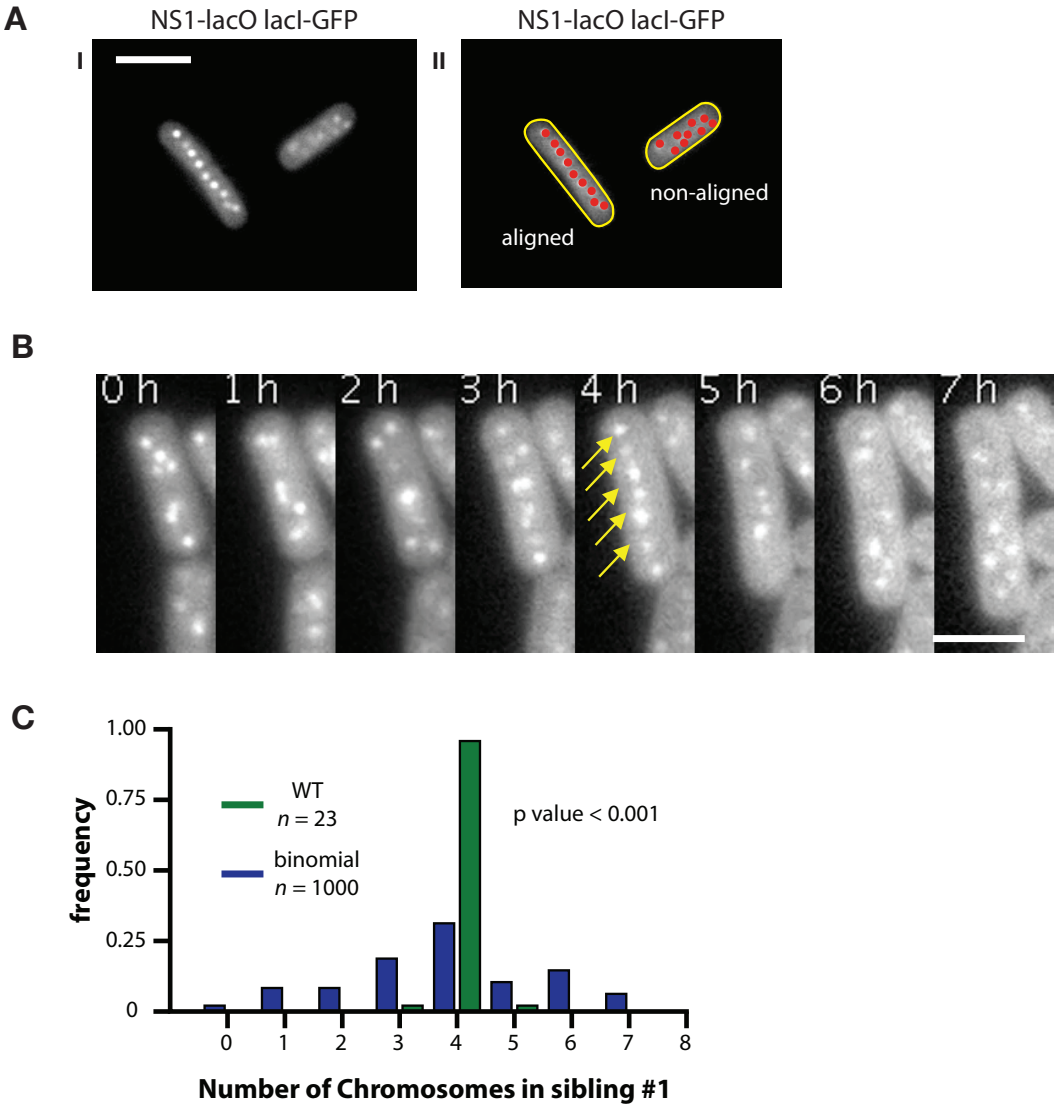

Figure 1.6: Chromosomes align transiently before non-random segregation occurs. (A) Chromosomes in cells visualized through LacI-GFP in an *NSI-lacO* background showed two different spatial arrangements. The chromosomes are either aligned along the major axis of the cell (left) or randomly localized (right). (B) Time-lapse imaging of single cells revealed a transient alignment of the chromosomes (4 hours, yellow arrows) approximately three hours before the cell entered cytokinesis. (C) Almost all sibling cells descended from mother cells containing eight chromosomes inherit four chromosomes, thus chromosome segregation is highly non-random.

1.4.7 Chromosomes are Non-randomly Segregated

The transient alignment of chromosomes prompted us to hypothesize that chromosome segregation is not random, and that the fidelity of chromosome segregation is high in cells harboring the transient alignment. This would be analogous to the spatial organization previously found regulating segregation of carboxysomes, microcompartments involved in the "carbon concentrating mechanism" in *S. elongatus*[143]. If true, chromosome segregation organization would be another example of order in bacteria, which were previously thought to be homogenous "bags of protein" with little to no internal organization[103]. To test our hypothesis, we assayed chromosome segregation by quantifying the number of chromosomes each daughter cell inherited after cytokinesis of mother cells (n = 48) that had 8 chromosome copies (Figure 1.6C). We found a striking difference between the experimental results and a predicted binomial distribution based on random segregation to daughter cells (Figure 1.6C Lilliefors test, P<0.001). A large majority of daughter cells inherited 4 chromosomes out of the 8 present in mother cells. Likewise, mother cells with 7, 9, and 10 chromosome copies were found to undergo nonrandom segregation to daughters (Lilliefors test, P<0.001).

1.5 Discussion

The goal of this research was to understand the dynamics of chromosome replication and segregation in *S. elongatus*. To do so, we used the GFP-LacI-*lacO* fluorescent repressor-operator system to visualize chromosomes *in vivo*. Similar analyses have been performed in other bacterial species, including *E. coli*, and *B. subtilis*[122,92]. Here, however, we investi-

gate live chromosomal dynamics in a bacterial species with multiple chromosomes. Having multiple copies of a single chromosome changes the parameters of the biological problem of replication and segregation that the organism must solve in order to successfully pass on its genetic information. The replication and segregation organization we found in *S. elongatus* differs from that in *E. coli* and other bacteria, and the difference can be attributed to the different challenges these bacteria face in passing on chromosomes to the next generation.

Bacterial chromosomal DNA is compacted into a nucleoid. In *E. coli*, the terminus and the origin can be overlapping or far apart from each other depending on the timing of replication and the cell cycle[151]. In our study, we labeled either the terminus or the origin of replication and analyzed origin movement in replication and segregation; however, different parts of the chromosome may be localized differently. We could not maintain the strain with the *lacO* array integrated at the terminus regions for long periods of time and we speculate that this may be due to a higher level of recombination occurring in that region of the chromosome. Integrations at other regions of the chromosome will clarify whether this trend is global or specific to that region of the chromosome, and whether the spatial and temporal patterns we observe can be generalized.

Our studies suggest that *S. elongatus* chromosome origins are randomly distributed throughout the cell. This is in contrast to specific localization of origins and termini in other bacterial species. In *E. coli* for example, the *oriC* is localized mid-cell and the terminus region is located at the poles of the cells[59]. In *C. crescentus*, both the *oriC* and *ter* are located at the poles[79]. In *V. cholerae*, one of its two origins is localized to the mid-cell, and the other is at the pole[46]. *B. subtilis* origins are located at the poles[173]. In these other species, replicated chromosomes initially occupy the same regions and require separation

for proper segregation to occur. The relative positions of *S. elongatus* origin and terminus are still unclear. Two color experiments with tagging at both the terminus and origin (or other regions of the chromosomes) would help elucidate details of replication and segregation: how the DNA strand is situated within the cell during different phases of replication and segregation. It will also answer how compact and region-excluded each chromosome is from its neighbors. The chromosomes may occupy distinct territories as has been previously observed in DAPI stained chromosomes[146]. The existence of local regions occupied by a single nucleoid can be advantageous for an organism with multiple chromosomes as it can undergo cytokinesis without the need to untangle chromosomes using FtsK or similar proteins.

Chromosome replication and cell division are fundamentally linked in most organisms[151], that is, replication must occur exactly once before (or during) cell division. Both events are believed to be regulated by mass doubling time, i.e. how long it takes for a cell to grow to double its mass[64]. Contrary to these studies in other organisms, asynchrony of chromosome replication has been observed in *S. elongatus*[8,172]. Indeed, recently, Watanabe *et al.* showed that cyanobacterial chromosomes replicate asynchronously based on mapping analysis and fixed cell staining methods[172]. They also suggested that while replication is still coupled to cell division (peaking at a few hours before cell division occurs), it is less stringently coupled than in *E. coli* or *B. subtilis*. In addition, cell division has been found to be gated by circadian rhythm[113], that is, cyanobacteria do not divide at night. Thus, timing of chromosome dynamics may be regulated by the circadian cycle. In the current study, we found that replication in *S. elongatus* is independent of cell division and is instead correlated to cell length. We also found that chromosomes were replicated multiple times every cell di-

vision. Both replication and cell division may still be dependent on cell mass, but if so, the threshold of activation is much lower for chromosome replication than it is for cell division in *S. elongatus*. Observing chromosome copy number in cells where cytokinesis is inhibited may uncouple these dependencies, allowing us to visualize only the correlation to cell mass independent of cell division. In addition, growing cells in multiple conditions leading to different cell division rates would also clarify the dependence between cell growth and chromosome copy number.

Several mechanisms ensure chromosomal replication takes place once every cell cycle and simultaneously from all origins in *E. coli* as well as in other prokaryotes[154,151]. These regulatory cellular processes give rise to a chromosome copy distribution where all possible values can be written in the form 2n, where n is an integer. In contrast, a Gaussian-like distribution of chromosome copy number similar to the one we observed has been found in *E. coli dnaA* mutants, which have disrupted initiation of replication[63,153]. Such a distribution could arise from the inability of some chromosomes to complete a single round of replication after the cell initiates synchronous replication. Alternatively, asynchronous initiation could also result in distributions containing copy numbers other than 2n. It is also possible that unequal partitioning could result in such a distribution. However, we found that most cells only have one replisome and therefore only one actively replicating chromosome at a time. We also found that partitioning was not unequal. These findings support asynchronous initiation as the mechanism that resulted in the observed Gaussian distribution of chromosome copy number.

There are two competing models for how replisomes proceed in DNA replication[5]. The first is the independent replisome model, in which replisomes and replication forks track

along the stationary chromosome. The second is the spooling replisome model, in which DNA is "spooled" through relatively stationary replisomes. A variation of the spooling model is the factory model in which the left and right replisomes are physically coupled throughout DNA replication. While early results favored a spooling model, recent results have shown that sister replisomes transiently separate. Another study observed that sister replisomes appear together early in S-phase but afterwards independently track DNA until they meet again as they near the terminus region. While our results suggest the spooling model is more likely in *S. elongatus*, we cannot rule out the independent model given the resolution limits of wide-field microscopy, especially if chromosomes each occupy exclusive local cytoplasmic regions during replication.

The spatial organization of chromosome origins along the major axis of the cell was observed for short periods (<1 hr) of the cell cycle. This spatial organization of chromosomes is surprisingly similar to eukaryotic mitosis, in which chromosomes align during metaphase before migration to the poles. However, in our current study, an active mechanism maintaining the organization was missing. Most other mechanisms of prokaryotic spatial organization studied such as carboxysome localization in *S. elongatus*[143] as well as low-copy plasmid segregation in *E. coli*[139,54] are maintained constantly throughout the cell cycle. We speculate that *S. elongatus* may not require an active segregation mechanism throughout the entire cell cycle in part because entropic forces may not be sufficiently disruptive to the arrangement of chromosomes post alignment and before cytokinesis. That is, due to having multiple chromosomes, *S. elongatus* does not have a stringent segregation problem and may not need constant organization. The transient alignment may enrich for rather than actively impose even segregation.

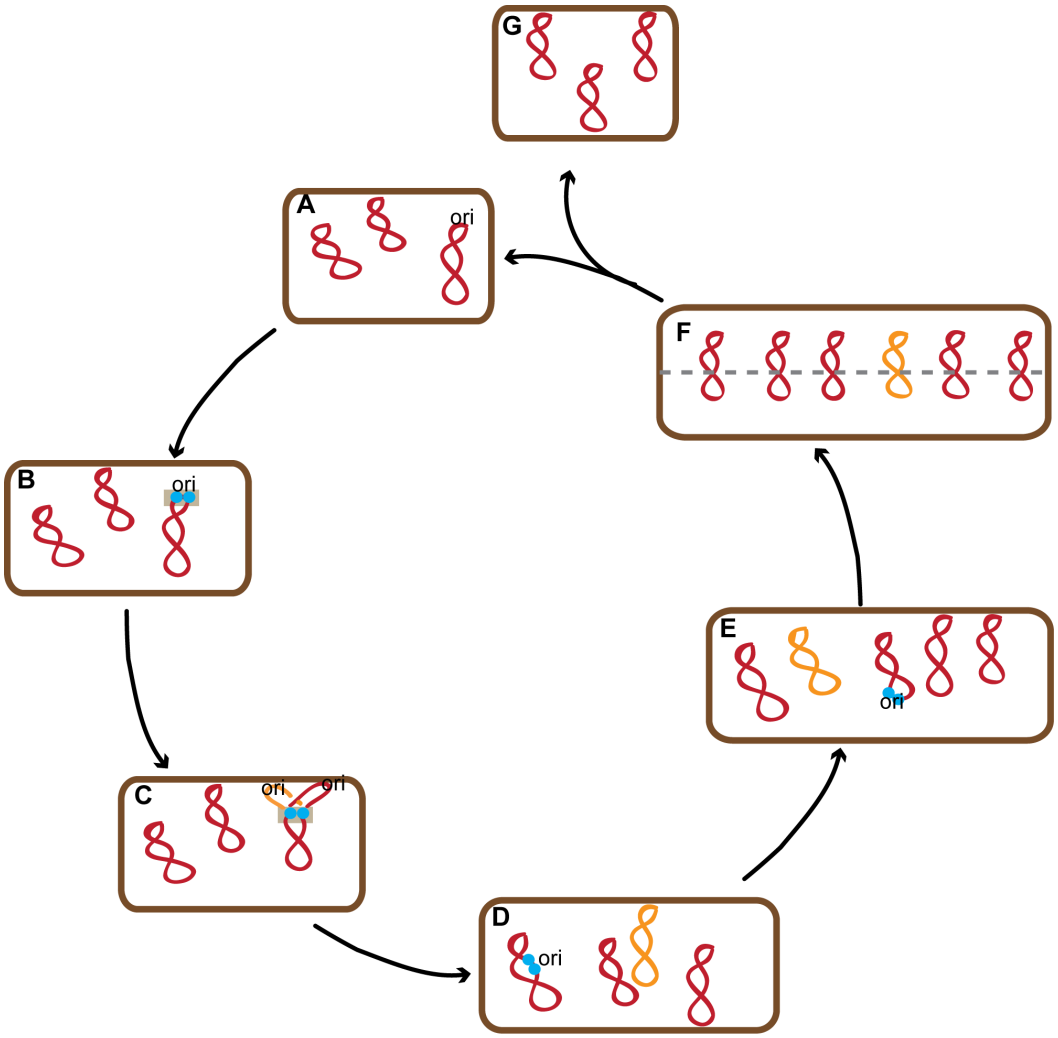

Figure 1.7: Model of chromosome replication and segregation in the polyploid bacterium *S. elongatus PCC 7942*.

S. elongatus possesses multiple copies of a single chromosome, shown in red (A). Chromosomes are duplicated asynchronously and coupled to cell growth (B, D, E). Newly synthesized chromosomes (orange) are synthesized in a spatially random manner (D,E,F). Replisomes (blue) assemble on a spatially random chromosome (B,D,E), but once initiated, their motion remains confined within the same region of the cell (grey box, B to C). Chromosomes transiently align (F) before non-random segregation and cytokinesis (F to G & A).

1.6 Conclusion

We have shown that the multiple copies of the chromosomes in *S. elongatus* can be tagged and tracked in living cells. Our model and findings are summarized in Figure 1.7. Chromosomes are replicated in a linear-like fashion correlated with cell length in growing cells. By tracking replisomes *in vivo* we show that chromosome replication takes place in a confined region of the cytoplasm in accordance with the spooling replisome model and that chromosome replication does not happen preferentially in specific locations of the cell. Finally, we show *S. elongatus* segregates chromosomes to daughter cells in a non-random fashion, which we speculate may be the result of a cellular process that transiently organizes the chromosomes just before completion of cell division.

2

The Bacterial Carbon-Fixing Organelle Is Formed by Shell Envelopment of Preassembled Cargo

2.1 ABSTRACT*

Cyanobacteria play a significant role in the global carbon cycle. In *Synechococcus elongatus*, the carbon-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) is concentrated into polyhedral, proteinaceous compartments called carboxysomes. Using live cell fluorescence microscopy, we show that carboxysomes are first detected as small seeds of RuBisCO that colocalize with existing carboxysomes. These seeds contain little or no

*Portions of this chapter are reproduced from the following work: Chen, A. H., Robinson-Mosher, A., Savage, D. F., Silver, P. A., & Polka, J. K. (2013). The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. *PLOS ONE*, 8(9), e76127. doi:10.1371/journal.pone.0076127. Reprinted with permission under the Creative Commons Attribution (CC BY) license.

shell protein, but increase in RuBisCO content over several hours, during which time they are exposed to the solvent. The maturing seed is then enclosed by shell proteins, a rapid process that seals RuBisCO from the cytosol to establish a distinct, solvent-protected microenvironment that is oxidizing relative to the cytosol. These closure events can be spatially and temporally coincident with the appearance of a nascent daughter RuBisCO seed. Carboxysomes assemble in a stepwise fashion, inside-to-outside, revealing that cargo is the principle organizer of this compartment's biogenesis. Our observations of the spatial relationship of seeds to previously formed carboxysomes lead us to propose a model for carboxysome replication via sequential fission, polymerization, and encapsulation of their internal cargo.

2.2 Introduction

Intracellular compartmentalization has long been considered the exclusive province of eukaryotes. However, prokaryotic cells also contain intracellular organelles, falling broadly into two categories. Some compartments are membrane-bound, including Gemmata nucleoids[49], cyanobacterial thylakoids[115], and magnetosomes[116]. Others are completely proteinaceous, such as gas vesicles[132] and metabolically active structures termed bacterial microcompartments. These form icosahedral structures that enclose enzymes required for certain metabolic processes, such as ethanolamine and propanediol utilization[83,183].

The carboxysome is one such microcompartment that encapsulates the carbon-fixing enzyme ribulose-1,5-bisphosphate carboxylase (RuBisCO) and carbonic anhydrase[152]. Carboxysomes are found in diverse cyanobacteria and chemoautotrophs and are crucial to the carbon sequestering capabilities of these organisms[22,12]. Inside the carboxysome, carbonic

anhydrase converts bicarbonate to CO2, which, along with ribulose-1,5-bisphosphate, is consumed by RuBisCO to produce 3-phosphoglycerate. Thus, the carboxysome serves to concentrate the metabolically inefficient RuBisCO enzyme, to increase the local concentration of CO_2, and possibly to decrease the local concentration of O_2. It has also been proposed that the carboxysome shell is selectively permeable to bicarbonate and ribulose-1,5-bisphophate while excluding oxygen, a competitor substrate of RuBisCO (Kinney et al. 2012). Finally, it has been speculated that the mature carboxysome must maintain a distinct internal oxidative microenvironment to enable the enzymatic activity of carbonic anhydrase[129].

The mechanism and temporal sequence of carboxysome assembly is not known. The interior of the carboxysome is densely packed with its major cargo RuBisCO and a lower concentration of carbonic anhydrase. These are enclosed by proteins that form an icosahedral shell about 100nm in diameter[162,84]. Though their ultrastructural, but not phylogenetic, similarity to viral capsids may suggest that carboxyosomes assemble *de novo*, the mechanism of their biogenesis remains an unsolved problem[83].

There is evidence to suggest that shell proteins and cargo assemble together. Partially assembled carboxysomes have been observed by electron cryotomography, always containing both RuBisCO and shell proteins[73]. However, these data also argue that the cargo must have some intrinsic ability to self-assemble, as RuBisCO is seen to fill the inner layers of the nascent compartment. Indeed, *in vitro* evidence suggests that carboxysome contents can self-associate to form a structure without shell proteins[125]. Shell proteins of some carboxysomes can also independently assemble, forming empty microcompartments in the absence of cargo proteins[112].

The sequence by which these proteins assemble to form this complex organelle is not understood. We employ live cell fluorescence microscopy of *Synechcoccus elongatus PCC 7942* to monitor the dynamics of carboxysome assembly. We find that carboxysomes originate near, and in some cases using material from, preexisting carboxysomes. They are born as small foci of RubisCO, which then grow over a period of hours. Shell proteins colocalize to these foci hours later, abruptly assembling to enclose the compartment and establish a protected internal microenvironment.

2.3 Materials and Methods

2.3.1 Bacterial strains and growth conditions

A table of all relevant strains and plasmids is presented in Table A.1. All chemicals were obtained from Sigma-Aldrich unless otherwise noted (St. Louis, MO). The wild-type *Synechococcus elongatus PCC 7942* strain was acquired from the American Type Culture Collection (ATCC, Manassas, VA). *S. elongatus* cells were grown in solid BG11 medium with an illumination of 2000 lux at 30°C[2]. *S. elongatus* were transformed following standard protocols by washing with 10mM sodium chloride followed by incubation overnight in the dark with 100 ng of plasmid DNA, and subsequently plating on selective media[21]. Antibiotics were used at the following concentrations: kanamycin 10 μg/ml, spectinomycin 50 μg/ml, and chloramphenicol 10μg/ml. 25μM isopropyl β-D-1-thiogalactopyranoside (IPTG) induction was used for RbcL-GFP or RbcL-mOrange. 50μM IPTG was used to induce formation of bar carboxysomes and 1mM IPTG for RbcL-10GFP.

2.3.2 Plasmid construction

Cloning was done using Gibson assembly unless otherwise noted. IPTG inducible GFP strain (pDFS724) was obtained as previously described[143]. This neutral site 2 (NS2) plasmid contains a region with lacI and a promoter from pTRC99a followed by RbcL-sfGFP. sfGFP in pDFS724 was replaced by mOrange2 to obtain pAHC003. The shell protein fusion pAHC134 was obtained by modifying pDFS594s[143], replacing YFP with sfGFP. Two color strains were obtained by double transformation of pAHC003 and pAHC134. RbcL-SNAP fusion plasmid pAHC126 was obtained by replacing sfGFP in pDFS724 with SNAP tag. RbcL- roGFP fusion plasmid pAHC149 was constructed using restriction cloning at NheI and NotI. sfGFP in pDFS724 was replaced with roGFP1 (University of Oregon Remington Laboratory).

2.3.3 Image acquisition

Cells were plated onto BG11 + 2% agarose pads with IPTG as necessary and placed on a glass bottom dish (Part No. P35G-1.5-20-C, MatTek, Ashland, MA). The addition of 100µl of water around dish edges and a paraffin film seal permitted long-term imaging.

FRAP image acquisition was performed on a Nikon Ti inverted microscope (Nikon Instruments, Melville, NY) with a MicroPoint laser targeting system (Photonics Instruments, Saint Charles, IL) controlling a 100mW solid state 488nm laser for photobleaching. Imaging was performed using a 100x 1.4 numerical aperture objective, an EXFO XL-120 (Lumen Dynamics Group, Mississauga, Canada) fluorescence light source, and an ORCA-R2 charge coupled device camera (Hamamatsu Photonics, Hamamatsu, Japan).

As previously described[143], all other imaging was done using a Nikon TE-2000 micro-

scope with a 100x 1.4 numerical aperture objective, a Lumencor LED fluorescence illuminator, and an ORCA-ER (Hamamatsu Photonics, Hamamatsu, Japan) charge coupled device camera. Acquisition was controlled using a custom MATLAB script controlling μManager[36] and a network AC power controller (IP Power 9258T) for photosynthetic lighting. Images were processed and analyzed with ImageJ[144].

2.3.4 Image analysis

For carboxysome lineage mapping, we used uTracker[77] to identify and localize closely-spaced point spread functions. The coordinates of these particles were imported into TrackMate, a plugin for FIJI[144] for manual annotation of track splitting events. Plots of lineages were retraced into vector format for counting division events and measuring maternal age.

For FRAP quantitation, the photobleaching rate after background subtraction was approximated with a linear function, which was used to correct measurements of the bleached region. Intensity was normalized to the maximum (in frame 1).

For ratiometric imaging, we background subtracted both 410nm and 488nm images with a 50-pixel radius rolling ball. We then registered the images with translations measured from imaging fluorescent beads and divided the 488nm image by the 410nm after converting to 32 bit format for floating point operations. This image was then multiplied by a mask of carboxysomes we generated based on a thresholded, 10-pixel rolling ball radius background-subtracted 410nm image. The mean intensities of regions larger than 9 pixels^2 were quantitated with the "analyze particles" features of FIJI.

2.3.5 Pulse-chase SNAP dye labeling

RbcL-SNAP strains that were induced with 25μM IPTG for 12 to 24 hours were labeled with SNAP-Cell BG 505-Star (New England Biolabs, Ipswitch, MA) following manufacturer's instructions. Briefly, 1mL of cells were spun down and resuspended in 100 μL BG11 with 25μM IPTG and 5μM dye substrate. Labeling was done for 30 minutes in light. Cells were washed 3 times with BG11 and resuspended in BG11 with 25μM IPTG for 30 minutes in light. Cells were washed once more with BG11 and then transferred to an agarose pad for imaging. Either time-lapse imaging at 1 hour intervals or two endpoints 1 day apart were taken. Analysis was performed by manually counting and measuring foci intensity in ImageJ.

2.3.6 Western Blotting

Cells were lysed by sonication in 3% SDS lysis buffer, and proteins were separated on NuPAGE Novex 4-20% Tris-glycine gels (Life Technologies, Grand Island, NY). Transfer to a nitrocellulose membrane was performed using the iBlot Gel Transfer Device and iBlot Gel Transfer Stacks (Life technologies, Grand Island, NY). Subsequent blotting was done using the SNAP-ID Protein Detection System (EMD Millipore, Bellerica, MA) following manufacturer's instructions. Polyclonal anti-RuBisCO antibody (Agrisera, Prod. ID AS03 037) was used at a final dilution of 1:5000 and an HRP-conjugated goat anti-mouse antibody (Abcam, ab97265) was used at a final dilution of 1:5000. Peroxidase conjugates were detected using SuperSignal West Dura Extended Duration substrate (Thermo Scientific).

2.4 Results

2.4.1 Growing cells typically assemble one carboxysome at a time

Given existing structural evidence, we reasoned that solvent-accessible labeling strategies could be used to mark only RuBisCO inside carboxysomes in the process of forming. We constructed a strain of *S. elongatus* with SNAP labeled RuBisCO (RbcL-SNAP) expressed under an IPTG-inducible promoter (Figure 2.1A). After a 24 hour induction, we pulsed the cells with a fluorescent cell-permeable BG dye for 30 minutes and visualized them with fluorescence microscopy (Figure 2.1A-B). Among cells containing labeled foci, 88% had one, 10% had two, and 2% had three foci (Figure 2.1B, n=442). Thus, only a small subset of the average 3.7 carboxysomes per cell are labeled[143]. Since not all carboxysomes are labeled, this indicates that mature carboxysomes are impermeable to the BG dye. Thus, BG-labeled foci represent carboxysomes being assembled during the labeling pulse. We used time-lapse microscopy to determine the difference between cells that contained labeled foci and those that did not. We found that all growing and dividing cells contained labeled foci (n=223), and those without labeled foci did not grow or divide, indicating that quiescent cells are not generating new carboxysomes. The distribution of foci numbers in growing cells indicates that *S. elongatus* carboxysomes are assembled one at a time rather than in parallel, as has been observed by electron microscopy in other genera of cyanobacteria[73].

2.4.2 New carboxysome are born colocalized with preexisting carboxysomes

Using time-lapse microscopy, we visualized the biogenesis of carboxysomes in live *S. elongatus* cells with green fluorescent protein labeled RuBisCO (RbcL-GFP) expressed under

an IPTG-inducible promoter (Figure 2.1C). Exogenous expression of the fusion protein produces an additional 11% of wild-type levels of RbcL (Figure A.1) and does not restrict growth[143]. We observed that new carboxysomes are formed at the site of preexisting carboxysomes. At the beginning of each birth event, a preexisting focus of RuBisCO is sometimes seen to take on an asymmetric character (Figure 2.1D, white arrow). Subsequently, a dimmer daughter focus emerges from the brighter mother carboxysome (Video S1).

The majority of new carboxysomes are generated at the site of preexisting carboxsyomes. Early after induction of RbcL-GFP, many unlabeled carboxysomes are still present in the cell, limiting our ability to determine whether all new carboxysomes colocalize with preexisting ones or arise at unrelated locations in the cell. To address this, we induced RbcL-GFP for 24 hours to label several carboxysomes in each cell, and then used particle tracking to generate lineage maps of carboxysomes (Figure 2.4). At the initiation of imaging, 65 "original" carboxysomes were present, and over the course of 26.1 hours, 106 new carboxysomes were formed. Of these, only two could not be assigned to visible mothers.

Furthermore, carboxysome birth events are spatially ordered, preferentially occurring at the quarter positions along the long axis of the cell (Figure 2.1J). These data are consistent with previous findings that cells have a mean of 3.7 carboxysomes positioned equally along their length by ParA (Savage, 2010). After birth, however, the new daughter carboxysome frequently localizes near the cell pole (Figure 2.1A and 2.1D). Quantification of RbcL-GFP intensity reveals that birth events are highly asymmetric, with an average daughter–mother intensity ratio of 1:4 (n=141, Figure 2.1K). We also observe that some birth events are correlated with rapid motions of either mother or daughter or both (Video S1). Indeed, automated tracking of carboxysome velocities suggests that carboxysome velocity is variable,

Figure 2.1 ***(following page)*****: Carboxysomes are born one at a time at the site of preexisting carboxysomes.** (A) In pulse-chase labeling of RbcL-SNAP in live *S. elongatus* cells, actively assembling carboxysomes with solvent accessible RbcL-SNAP are labeled with BG dye. Red: phase contrast. Green: RbcL. Scale bar: 1μm. (B) The distribution of the number of SNAP labeled carboxysomes, indicating active assembly, in cells directly after labeling (n=442). (C) The biogenesis of carboxysomes can be monitored from long timelapses. Red: phase contrast. Green: RbcL-GFP. Scale bar: 1μm. (D) Montage showing the formation of new carboxysome at the site of a preexisting carboxysome. White arrow indicates the birth event. Panel height: 25 pixels. Time interval: 3 minutes. (E–F) RuBisCO foci elongate into bar carboxysomes that subsequently split into two carboxysomes. Scale bar: 1μm. Time interval: 75 minutes. (G–I) Kymographs of RbcL-GFP in growing and dividing cells. Carboxysome birth events are indicated by white arrows. Scale bar: 1μm. Time interval: 3 minutes. (J) Spatial distribution of 234 birth events along the long axis of the cell. Quarter cell positions are favored. (K) Relative intensity of 141 pairs of new (daughter) carboxysomes and the preexisting carboxysomes to which they initially colocalize (mothers) reveals that birth events are highly asymmetric, with mean daughter intensity being 1/4 that of the mother. Because pairs are sorted into dim (daughter) and bright (mother) pairs, no data points can fall into the shaded area. Dotted line indicates a 1:4 ratio.

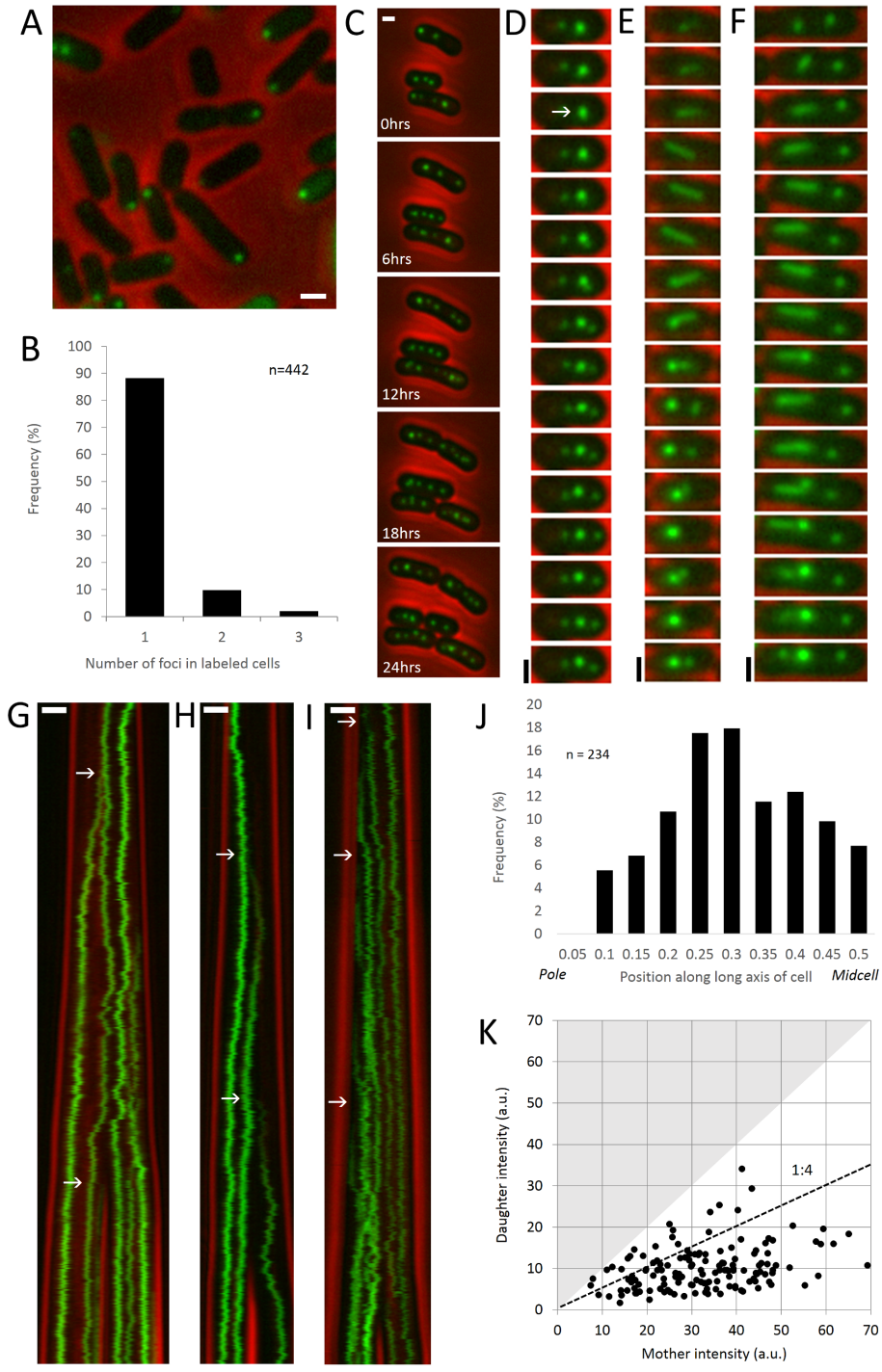

Figure 2.1: (continued)

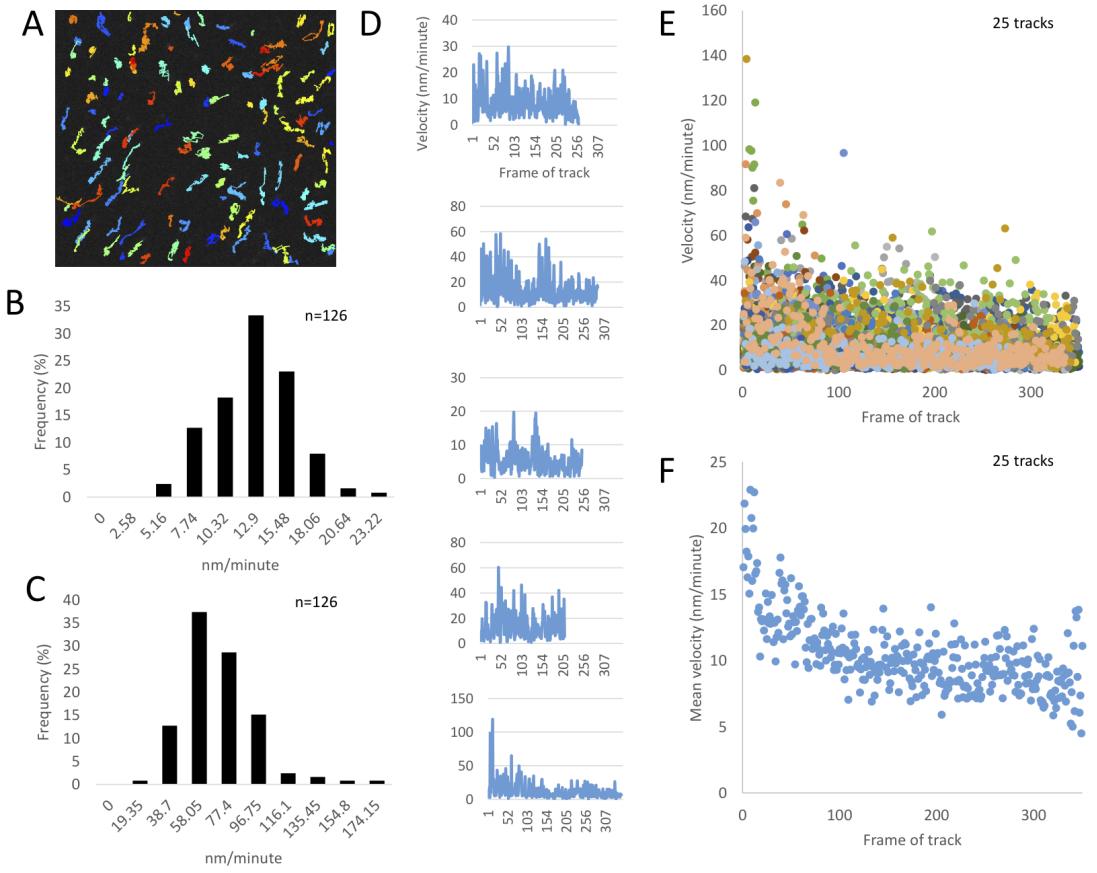

Figure 2.2: Mean and maximum velocities of carboxysome motion. 126 carboxysomes were tracked over a minimum of 85 frames, and the mean and maximum velocity of each track quantified. (A) The tracks overlaid on one frame of the movie. Different colors represent different tracks. (B) The mean of the mean carboxysome velocity is 11.5nm/minute. (C) The mean of the maximum carboxysome velocity is 60.6nm/minute. (D) Velocity is variable across each track, with a subset showing maximal velocity near the start of the track. (E) 25 tracks were selected at random from this set, and velocity was plotted against the frame number (ie, age) of each track. Interval of acquisition, 5 minutes. (F) The mean velocity per frame number across these 25 tracks. As track length is variable, fewer data points contribute to the mean toward higher frame numbers.

with the mean maximum speed of a carboxysome being over 60nm per minute (Figure 2.2C). This may be related to the approximately 100nm per minute movement of ParA, assuming a 3μm cell[143]. Analysis of individual tracks reveals that mean carboxysome velocity is higher in the first several hours after birth (Figure 2.2F).

Colocalization of birth events is especially obvious in our observations of bar carboxysomes. Bar carboxysomes are elongated carboxysomes (1-3μm in length) that provide a means to study the spatial organization of carboxysomes above the resolution limit of light microscopy. They are thought to be assembly intermediates or misassembled carboxysomes, especially since they are present in higher frequencies (up to 20% of all carboxysomes) when under environmental carbon limitation or in cells with mutations that compromise carbon fixation[125,111]. However, they are well documented by electron microscopy in normal *Synechococcus* cells[73,51] as well. In our RbcL-GFP strain 0.5% of labeled carboxysomes are bars after induction with 25μM IPTG (n=191 total), and 1% are bars after 50μM IPTG (n=395 total). These carboxysomes colocalize with shell protein, though their redox state suggests that they are immature (Figure 2.3, compare to Figures 2.5 and 2.6). Bar carboxysomes begin as puncta that elongate and subsequently collapse or split into two carboxysomes (Figure 2.1E-F and Video S1).

2.4.3 Maternal age influences the frequency of carboxysome births

By further analyzing the lineage maps, we found that new carboxysomes are more likely to be born near recently formed carboxysomes than near older ones. For each birth event, we measured the age of the mother carboxysome if it was born during the course of our observations (vertical dotted lines in Figure 2.4A and histogram in Figure 2.4C, black bars).

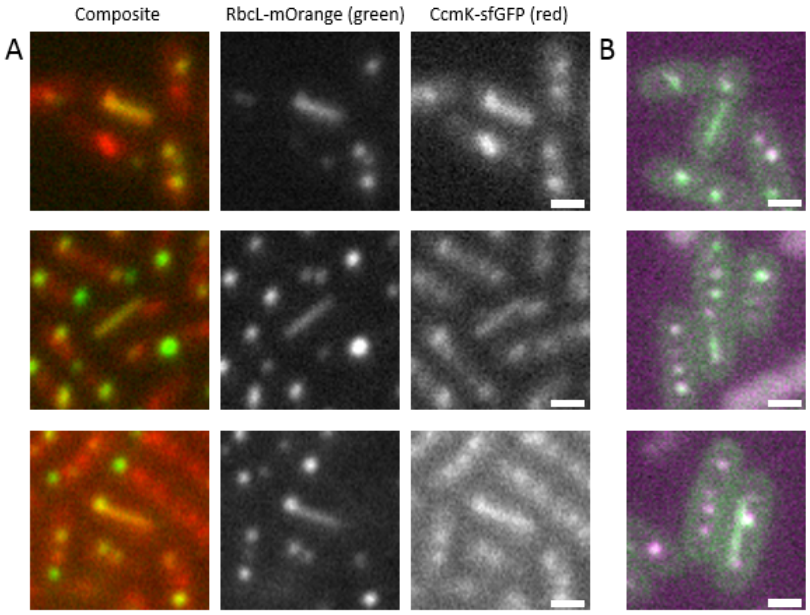

Figure 2.3: Bar carboxysomes colocalize with shell but are not oxidized. (A) Bar carboxysomes contain both RuBisCO and shell protein. Red, CcmK4-GFP. Green, RbcL-mOrange. Scale bar, 1µm. (B) Bar carboxysomes are relatively reduced compared to punctate carboxysomes. Still frame composite images of 488nm (reducing, green) and 408nm (oxidizing, purple) RbcL-roGFP1 as in Figure 2.6. Scale bar, 1µm.

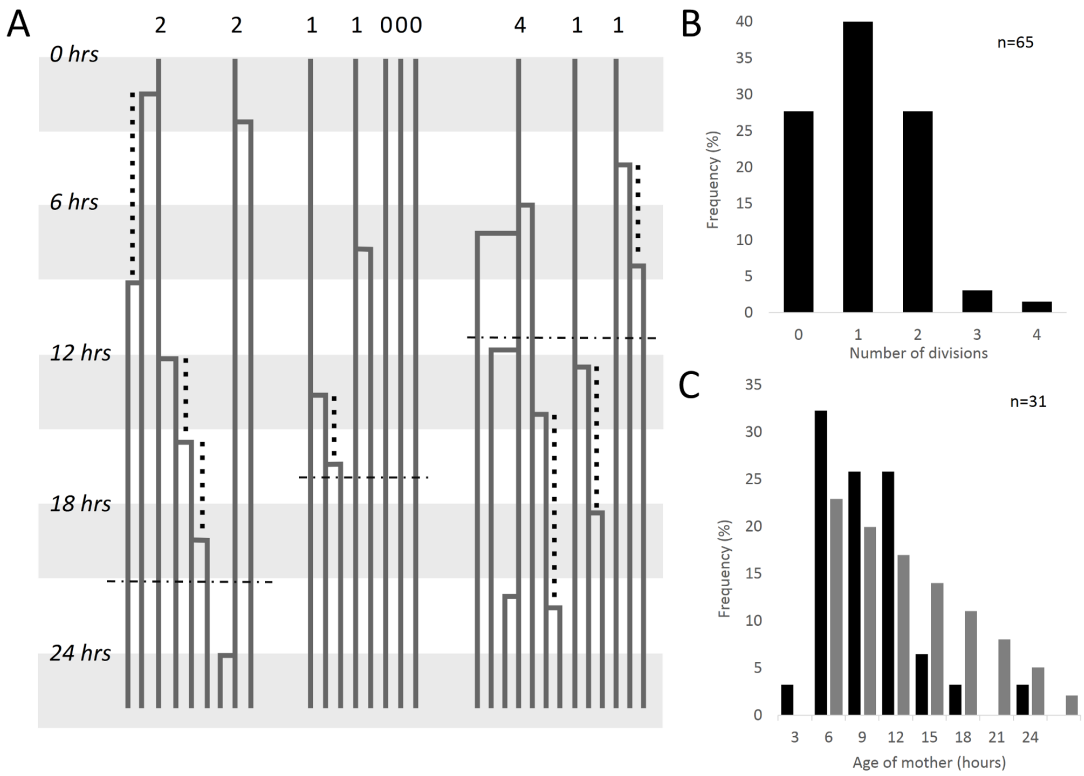

Figure 2.4: Mapping of carboxysome lineages reveals that new organelles undergo an initial refractory period before producing daughters of their own.(A) Example lineages of carboxysomes from three out of 25 total cells analyzed from a 522 frame movie taken at 3 minute intervals over approximately 26.1 hours. Each vertical solid line represents a carboxysome tracked through time, with right-angle connectors joining daughters to mothers. Digits at the top of the panel indicate the number of times carboxysomes present at the beginning of the movie have colocalized birth events over the course of the analysis (26.1 hours), represented in the histogram in panel B. Vertical dotted lines indicate the length of time between the birth of a mother and the colocalized birth event of a daughter to that mother, i.e. the measurable age of the mother when a daughter appears. The ages are represented in the histogram in panel C. Horizontal dotted lines indicate the time of cell division. (B) Histogram of the number of births colocalized to original carboxysomes in the entire dataset (n = 65). (C) Histogram of measurable ages of mothers tabulated over the entire dataset (n = 31). Black bars indicate observed distribution of birth ages, while grey bars indicate the theoretical distribution predicted by an age-independent model. The two distributions are statistically significantly different (KS-test, p=0.04).

Strikingly, after a carboxysome is born, there is a marked refractory period until a new colocalizing daughter appears. This is characterized by a lack of births in the first three hours of its lifetime and suggests the structure must mature before another birth event occurs. Immediately following this refractory period, there is a burst of birth events. However, our imaging interval favors the observation of early birth events over late ones. To determine whether this apparent burst is significant, we compared our data to a model where birth probability is constant regardless of carboxysome age, adjusting for the limitations of our imaging interval. Our observed distribution of birth ages (black bars, Figure 2.4C) is significantly different from the theoretical distribution predicted by the age-independent model (grey bars, Figure 2.4C) (Kolmogorov-Smirnov test, h=1, p-value = 0.0416, k=0.2438). In comparison, our data show that birth rates are enriched in the first 12 hours of the lifetime of the mother. We measured the birth rate in the 9 hours following this three hour refractory period at 0.42 per carboxysome (n=57 carboxysomes), versus 0.28 per carboxysome over 9 hours for those at least 12 hours old (n=65).

While young carboxysomes have colocalized birth events more frequently during the burst period, mature carboxysomes do not show preference in the timing of colocalized birth events—they have daughters randomly. By tabulating the number of times that carboxysomes visible at the start of the time lapse colocalized with new birth events, we found that this distribution of events is nearly Poissonian, with a mean of 1.1 and a variance of 0.8 (Figure 2.4B). This indicates that births near mature mothers is a random process. Furthermore, the probability that a preexisting carboxysome has daughters in the second half of the movie is not influenced by whether it did (0.36) or did not (0.37) in the first half of the movie, suggesting that births near mature carboxysomes are independent events. Further-

more, timing of carboxysome birth events is not correlated to cell divisions (horizontal dotted lines in Figure 2.4A). The distribution of birth event timing as a fraction of the cell cycle is not significantly different from a random uniform distribution across the cell cycle (Kolmogorov-Smirnov test, h=0, p=0.3222, kstat=0.13).

2.4.4 Carboxysome biogenesis begins with a sigmoidal assembly of RuBisCO

To understand the nature of the maturation process, we followed the assembly of RuBisCO and shell protein (CcmK4). We first examined the kinetics of RuBisCO assembly by measuring the intensity of RbcL-GFP foci over time. This indicated that RuBisCO assembles over the course of many hours in distinct phases that display sigmoidal kinetics (Figure 2.5A). While individual carboxysomes assemble at different rates, we observed three regimes: a lag phase, followed by rapid assembly, and finally a plateau phase - assembly kinetics reminiscent of nucleation condensation polymers.

Once assembled, RuBisCO does not freely diffuse inside the carboxysome. To probe the nature of assembled RuBisCO-GFP, we monitored fluorescence recovery after photobleaching (FRAP) of bar carboxysomes (Figure 2.5B-C and Figure A.2). Bar carboxysomes are sufficiently large such that only a segment of the bar was bleached (Figure 2.5B solid box) while the rest of the bar remained fluorescent (Figure 2.5B dashed box). No recovery of fluorescence was seen up to 150 seconds after bleaching (Figure 2.5C). RuBisCO hexadecamers are roughly 500kDa in size, and freely diffusing protein complexes of similar molecular weight have been reported to recover in *in vivo* FRAP experiments in less than two seconds[57]. This discrepancy indicates that assembled RuBisCO does not freely exchange with monomers in the cytoplasm or in the rest of the carboxysome; rather, assembled cargo

Figure 2.5 ***(following page)*****: RuBisCO slowly forms a structured assembly prior to rapid colocalization of shell protein.** (A) RuBisCO assembly, as measured by fluorescence intensity, follows sigmoidal kinetics. Each trace represents a new carboxysome. Cell is the same as that depicted in Figure 2.1I. Imaging interval: 3 minutes. (B) Fluorescence recovery after photobleaching of a segment of a bar carboxysome. Solid box shows bleached area. Unbleached area (dashed box) was used for photobleaching correction. Cells were imaged at regular intervals after bleaching to assay for recovery. Scale bar: 1µm. (C) Quantification of FRAP in (B). Grey bar indicates bleaching event, when fluorescence sharply decreases. No recovery was seen after 150 seconds. (D–H) Time lapse of RbcL-mOrange (green) and CcmK4-GFP (red). Arrows indicate birth events of carboxysomes. Newly born RuBisCO initially buds off without shell protein. Shell protein colocalizes to RbcL-GFP foci hours after birth. In some cases (G and H), shell protein assembly is correlated with the formation of a new RuBisCO focus. Scale bar: 1µm. Time interval: 25 minutes. (I–K) Kymograph of RbcL-mOrange (J) and ccmK4-GFP (K) assembly. Shell protein assembly (yellow arrow in K) initiates well after RuBisCO birth event (yellow arrow in J). Scale bar: 1µm. Time interval: 5 minutes. (L) Individual trace of the fluorescence intensity of a CcmK4 focus in the process of formation. Time interval: 5 minutes.

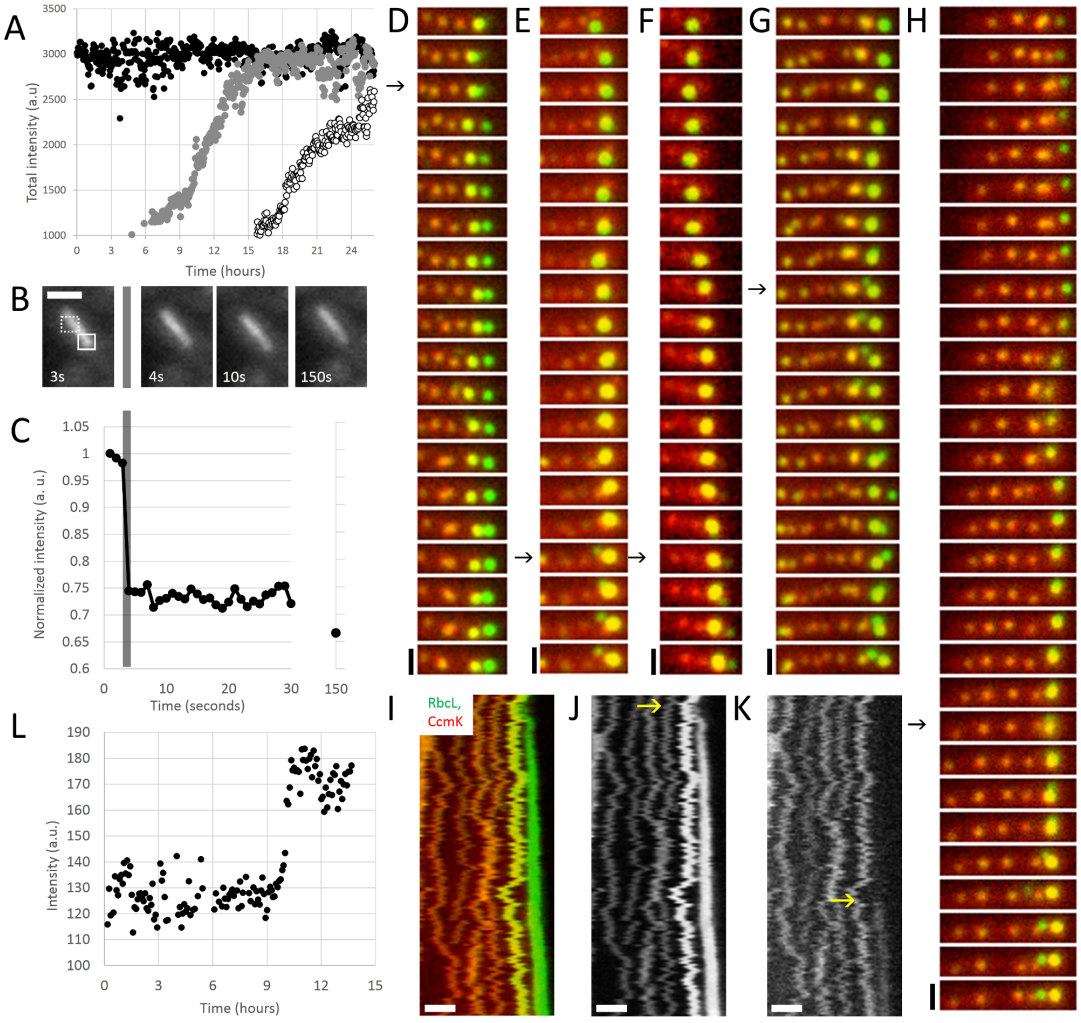

Figure 2.5: (continued)

is static on the timescale of minutes.

2.4.5 Shell proteins rapidly colocalize with RuBisCO late in the assembly process

To determine the relative kinetics of RuBisCO and shell assembly, we performed time-lapse microscopy of cells expressing (1) inducible RuBisCO fused to mOrange, and (2) the shell protein CcmK4 fused to GFP under a constitutive promoter. At the beginning of the observation interval, recently born carboxysomes show strong signal in the RuBisCO channel, while old carboxysomes show weak, background levels of signal. New carboxysomes begin assembly with little to no detectable shell protein (Figure 2.5D–H and Video S2). Instead, shell protein associates with nascent carboxysomes at a later point: the mean time between the first appearance of RuBisCO and detectable shell protein was 4.7 hours (+/-2.2 hours, n=54). Kymographs of the formation process are shown in Figure 2.5I–K, where the shell protein suddenly colocalizes 8 hours after the birth event of the RbcL-GFP focus. In contrast to the many hours required for RuBisCO assembly (Figure 2.5A), the assembly of shell protein completes rapidly; shell intensity reaches steady state in less than two hours (Figure 2.5L and Figure A.3). Interestingly, in a fraction of cases, we observed the apparent birth of small daughter focus from a carboxysome 3.1 hours ($\pm$ 1.1 hours, n=9) after detectable shell colocalized with the mother (Figure 2.5D–H). The timing of these events correlates with the burst of births following a maternal refractory period in lineage maps (Figure 2.4C).

2.4.6 The carboxysome establishes a unique microenvironment late in the assembly process

The enzymatic activity of carbonic anhydrase relies on an environment more oxidative than the bacterial cytosol[129]. This predicts that the carboxysome must maintain an internal oxidizing state. To monitor changes in the redox state of the carboxysome over time, we tagged RuBisCO with the redox-sensitive roGFP1[66]. The excitation spectrum of this protein shifts from one dominated by a maxima at 488nm under reducing conditions to one dominated by a maxima at 410nm under oxidizing conditions. By measuring the ratio of these two channels, we observed that carboxysomes display varying redox states within the same cell (Figure 2.6A). Over the entire population, we find that carboxysomes are distributed across a range of redox states (Figure 2.6B) but that the distribution is skewed toward oxidizing states (median=440, mean=487).

The late assembly of shell proteins on nascent carboxysomes predicts that maturing foci of RuBisCO share the reducing cytosolic environment. Indeed, monitoring changes in redox state revealed that newly formed RuBisCO foci are relatively reduced, regardless of whether imaging is started 6 (Figure 2.6C-D) or 24 (Figure 2.6E–G) hours after induction. As the carboxysome matures, RuBisCO-roGFP1 oxidizes, indicating the establishment of a distinct microenvironment (Video S_3). Though the low signal from roGFP1 prohibits the time resolution required to measure the carboxysome lineage, the most recently synthesized carboxysome typically oxidizes before a new one appears (Figure 2.6C–F). Rarely, a carboxysome is apparently born from a mother rapidly switching between oxidized and reduced states (Figure 2.6G). Taken together, our data show that new carboxysomes are born concomitant with shell closure and establishment of the oxidizing microenvironment in the

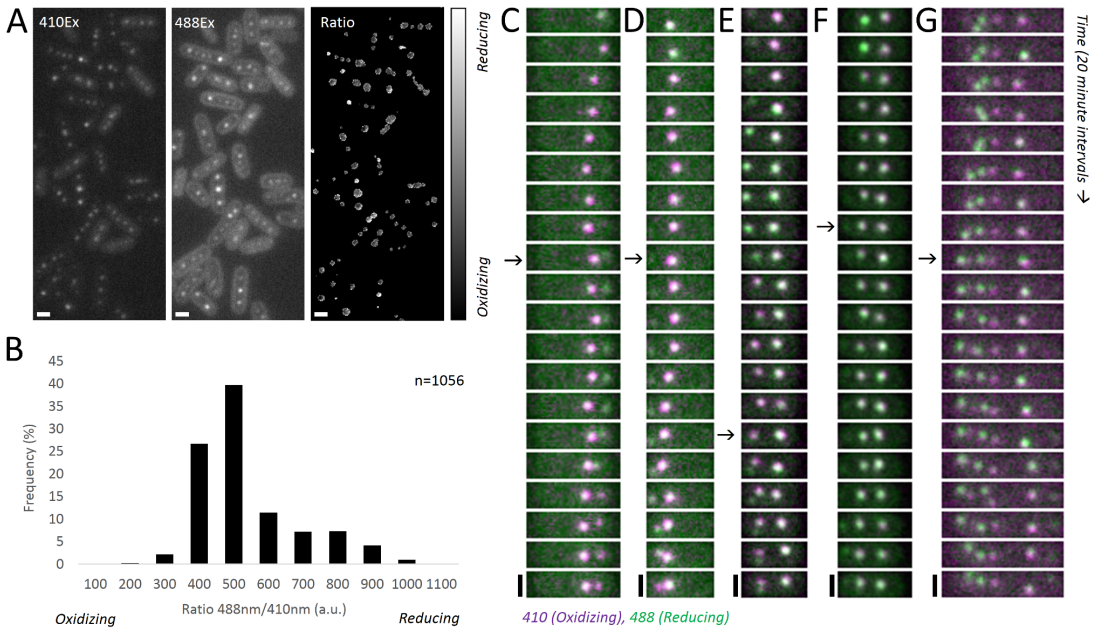

Figure 2.6: The carboxysome oxidizes over the course of its maturation. (A) RbcL-roGFP1 excited with 410nm (left) and 488nm (middle) produces ratiometric (488nm/410nm) differences in emission (right). Scale bar: 1µm. (B) A histogram of this ratio measured at each carboxysome focus reveals an asymmetric distribution biased toward a relatively oxidized state. (C–F) Montages of RbcL-roGFP1 show transitions from predominantly 488nm excitation (green) to 410nm excitation (magenta) over the maturation period of carboxysomes. Carboxysomes establish an oxidizing state before the appearance of a new carboxysome, rarely reopening to the cytosol after an initial closure (G). Arrows indicate birth events. Scale bar: 1µm. Interval: 20 minutes.

synthesized carboxysome.

2.4.7 Carboxysomes persist over the cell cycle and their cargoes can be redistributed to daughter carboxysomes

We used the BG pulse-chase experiments (Figure 2.1A-B) to track the lifetime and fate of RuBisCO assemblies over days. We observed that labeled RuBisCO from one initial focus can partition into two or more daughter carboxysomes, and that it persists over the time

interval of the experiment (45hrs) (Figure 2.7A). In some cases, the intensity of the mother carboxysome could be seen to decrease with the birth of a new focus (Figure 2.7B), indicating repartitioning of RuBisCO to new carboxysomes from old ones. In other cases, no decrease in mother intensity was detectable with the appearance of other BG foci. This is perhaps due to either splitting events with signal changes beneath our detection limit or the assembly of residual labeled cytosolic RuBisCO (Figure 2.7B, grey trace). During the period of this pulse-chase experiment, the number of labeled foci per cell was either invariant or increased (Figure 2.7C). Interestingly, disappearance of foci was never observed, suggesting that carboxysomes are not degraded, but rather passed down through generations.

2.5 Discussion

We show that the *in vivo* biogenesis of carboxysomes occurs by preferential assembly on preexisting RuBisCO structures that later separate from mother carboxysomes. These stable, cytosol-accessible nuclei grow over a period of hours until shell proteins abruptly enclose the carboxysome, establishing a microenvironment distinct from the cytosol. This maturation can be coincident with the release of a new colocalizing daughter seed of RuBisCO.

Carboxysomes are the major carbon-fixing centers of the photosynthetic cyanobacterium *S. elongatus*; thus, maintaining an appropriate number of organelles is vital to the cell[143]. The assembly of these compartments is regulated in two ways: 1) by formation of one carboxysome at a time, and 2) by regulation of their geometry. Our data suggest that both of these constraints arise from the assembly properties of the components.

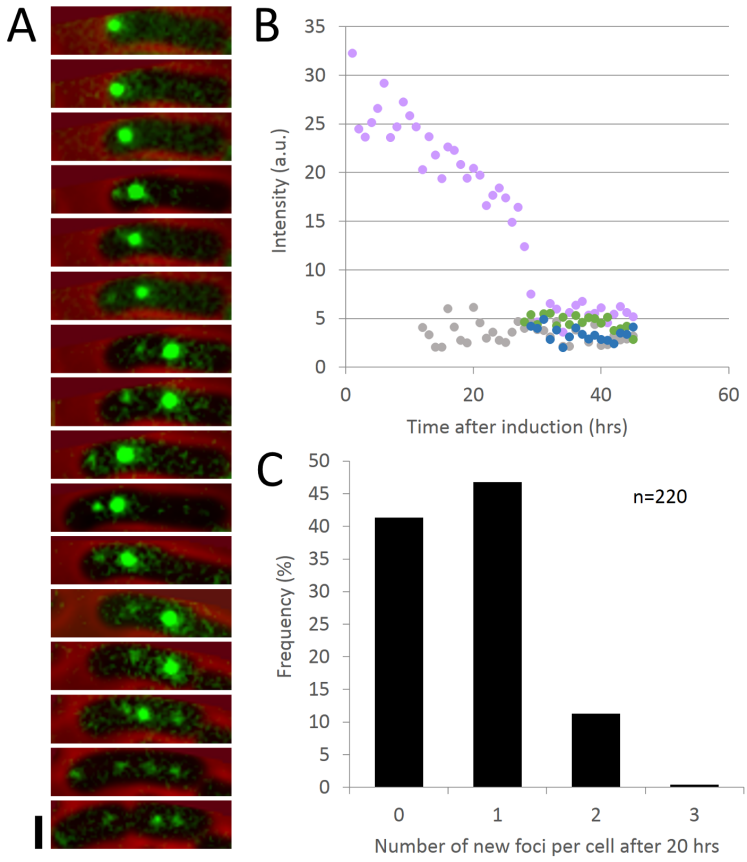

Figure 2.7: A solvent-accessible dye pulse labels foci that subsequently divide, but do not dissipate. (A) Montage showing one labeled RuBisCO focus partitioning into two or more daughter carboxysomes and persisting over the time interval of the experiment. Red: phase contrast. Green: RbcL-SNAP. Scale bar: 1μm. Time interval: 2 hours. (B) Intensity of the mother carboxysome (pink trace) sometimes decreases when new daughters (green and blue traces) are born. In other cases, the decrease is not detectable (grey trace). (C) Distribution of the number of new carboxysome foci formed per cell over the course of an experiment (20 hours). RuBisCO foci either persisted or divided over 20 hours (n = 220). All original foci were detectable at the end of the experiment.

2.5.1 Cargo assembly is the primary organizer of carboxysome biogenesis

Carboxysome biogenesis is tuned to produce one structure at a time (Figure 2.1B). This may be an energetically efficient strategy, as focused assembly minimizes the net time carboxysomes spend in an incomplete state, during which they cannot deliver energetic benefits to the cell. In order to achieve focused assembly, there must be a kinetic barrier to spontaneous nucleation of cargo so that growth occurs only on preformed seeds.

Several lines of evidence support a nucleation-limited assembly mechanism of RuBisCO. First, the one-at-a-time assembly process suggests that templated assembly is favored over *de novo* nucleation. Second, the sigmoidal kinetics of RuBisCO assembly are reminiscent of nucleation-limited polymers. Third, the elongation of some RuBisCO seeds into bar carboxysomes suggests that this assembly is an extensible process, capable of producing structures far larger than mature icosahedral carboxysomes. Fourth, FRAP of bar carboxysomes demonstrates that assembled RuBisCO does not freely diffuse, again reminiscent of a polymer lattice with stabilizing interactions between neighboring subunits.

Previous work also supports this idea. Contents from purified carboxysomes can self-assemble in a concentration-dependent manner *in vitro*[125]. Furthermore, RuBisCO inside carboxysomes is organized into a lattice[73,72,145] implying that multiple self-associating interactions direct cargo to fill the interior layers of the carboxysome.

In addition, it would be interesting to explore the localization patterns of the other cargo of carboxysomes, carbonic anhydrase (CcaA and CcmM), and determine its role in carboxysome nucleation. It has been suggested that CcaA is a part of the "scaffolding complex" for RuBisCO[135] and that CcmM has two isoforms and forms direct contacts with the shell and with RuBisCO[102]. Fluorescent tagging of these proteins and visualization could

provide further evidence on the role of the cargo in carboxysome biogenesis. However, as there are only 80 copies/carboxysome (compared to 2000 copies of RuBisCO per carboxysome)[68], visualization of carbonic anhydrase may pose challenges requiring bright fluorescent proteins or single molecule microscopy techniques.

2.5.2 Shell assembly specifies organelle size and limits further addition

Polymerized cargo appears to be stable and capable of extending far beyond the geometry of a mature icosahedral carboxysome, as suggested by the existence of bar carboxysomes. However, most mature carboxysomes are homogeneous in size. We propose that the rapid enclosure by the shell protein not only limits further cargo assembly by isolating the assembled RuBisCO from the cytosolic pool of subunits, but also sets the size of the carboxysome. Our data suggest two mechanisms for the size determination of RuBisCO assemblies: 1) size-selective enclosure, and 2) bisection of excess cargo.

We observe the assembly of shell protein only late in the biogenesis process (Figure 2.5I–K), presumably when the RuBisCO lattice reaches a given size. The topology of the growing RuBisCO seed thus would present a multivalent binding surface, with curvature depending on the size of the overall assembly. It is known that shell protein also self-associates into structures of a given radius[112]. Therefore, we speculate that when the curvature of the RuBisCO assembly matches that of the shell, RuBisCO-shell interactions organize shell-shell interactions, facilitating the assembly process. In other words, the intrinsic structure of the shell may act as a topological sensor that regulates timing of RuBisCO enclosure, ensuring that nascent carboxysomes reach a minimum size before encapsulation.

We also speculate that the polymerization of the shell can bisect a RuBisCO assembly

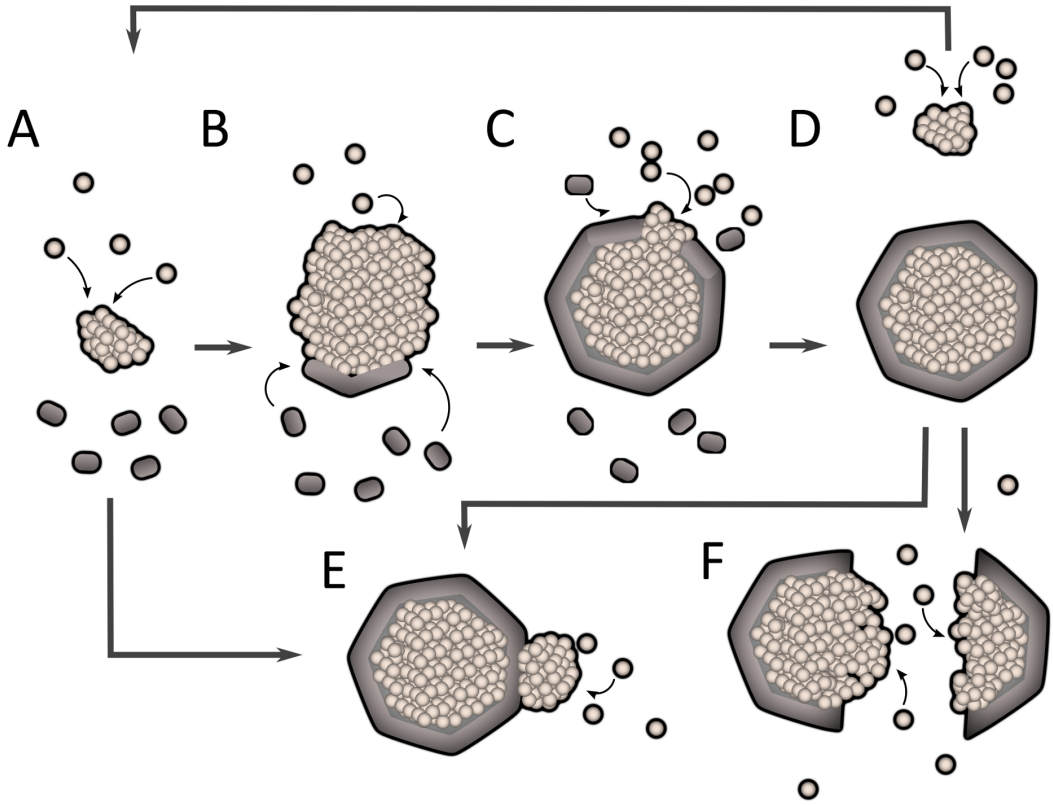

Figure 2.8: Model of carboxysome assembly. (A) RuBisCO seeds assemble from protomers over time. (B) Late in the assembly process, shell proteins rapidly assemble around RuBisCO. (C) Shell closure completes the carboxysome to establish an oxidizing environment, sealing RuBisCO from the cytosol. (D) A new RuBisCO nucleus forms after completion of the previous carboxysome. Colocalization may be driven by bisection of excess cargo by shell closure, or (E) by affinity of RuBisCO assemblies initiated elsewhere to the outside of the shell. (F) Rupture of a complete carboxysome would expose old RuBisCO cargo to template new assembly.

to generate a mature carboxysome and a new RuBisCo seed (Figure 2.8C and D). This shell-mediated pinching hypothesis presents a parsimonious explanation for the increased birth rates from young mothers (Figure 2.4) and the coincidence of new seed formation with both shell association and the establishment of a distinct microenvironment in the mother (Figures 2.5 and 2.6). Our data also support non-pinching mechanisms of templated carboxysome replication. For example, our pulse-chase data indicates that mature carboxysomes can fracture, as we observe repartitioning of RuBisCO to two daughter carboxysomes (Figure 2.8F).

While our data support some forms of replicative biogenesis, our methods cannot discriminate between *de novo* and templated nucleation events. It is plausible that RuBisCO seeds assembled *de novo* may be brought into close proximity with preexisting carboxysomes by other mechanisms. The carboxysome itself may be sufficient to capture independent seeds: crystal packing evidence from other studies suggests that shell proteins may contact one another face-to-face or assemble into antiparallel strips[162,11]. This may expose cargo-interacting surfaces to the outside of the carboxysome, creating affinity for cargo on the exterior as well as the interior carboxysome surface (Figure 2.8E).

In summary, all proposed mechanisms rely on the self-association of RuBisCO as the primary organizer and driving force of carboxysome biogenesis, with shell protein defining organelle geometry.

2.5.3 Broader implications

In addition to being crucial for global carbon fixation, the carboxysome has been proposed as a potential protein nano-factory capable of compartmentalizing heterologous reactions

for metabolic engineering purposes[9]. An N-terminal peptide has been identified for the targeting of cargoes to 1,2-propanedeiol utilization microcompartments[44], but such a mechanism in carboxysomes has been elusive. Our studies of an assembly process dependent on self-association of cargo and the establishment of a unique internal microenvironment will inform the design of any future systems.

3

Transplantability of a Circadian Clock to a Non-Circadian Organism

3.1 ABSTRACT*

Circadian oscillators are post-translationally regulated and affect gene expression in autotrophic cyanobacteria. Oscillations are controlled by phosphorylation of the KaiC protein, which is modulated by KaiA and KaiB. However, it remains unclear how time information is transmitted to transcriptional output. Here, we show reconstruction of the KaiABC oscillator in the non-circadian bacterium, *Escherichia coli*. This orthogonal system shows circadian oscillations in KaiC phosphorylation and in a synthetic transcriptional

*Portions of this chapter are reproduced from the following work: Chen, A. H., Lubkowicz, D., Yeong, V., Chang R. L., Silver P. A. (2015). Transplantability of a Circadian Clock to a Non-Circadian Organism. *Science Advances*. 05 June 2015: Vol. 1 no. 5. Reprinted with permission from AAAS.

reporter. Co-expression of KaiABC with additional native cyanobacterial components demonstrates a minimally sufficient set of proteins for transcriptional output from a native cyanobacterial promoter in *E. coli*. Taken together, these results demonstrate that a circadian oscillator can be transplanted to a heterologous organism for reductive study as well as wide-ranging applications.

3.2 Introduction

Circadian oscillators regulate metabolic and behavioral changes in anticipation of day-night variations in environmental conditions and are shared by organisms across the tree of life[37]. In the cyanobacterium *Synechococcus elongatus*, a well-studied system for circadian rhythm regulation, KaiC phosphorylation state changes are coordinated with diurnal cycles. During the day, KaiA promotes the autokinase activity of KaiC, resulting in phosphorylated KaiC at subjective dusk. At night, KaiB antagonizes the activity of KaiA, promoting the autophosphatase activity of KaiC, returning KaiC to the unphosphorylated state again at subjective dawn (Figure 3.1A)[85,167]. These changes drive corresponding oscillations in gene expression. The KaiABC system has been reconstituted *in vitro* and phosphorylation of KaiC oscillates in this minimal system[119]. These findings suggest that the cyanobacterial oscillator could function when heterologously expressed in other organisms, which would have practical applications as well as implications for understanding oscillatory behavior evolution. Here we show function of the reconstructed kai circadian oscillator in *E. coli*, a gut bacterium without its own native circadian rhythm.

3.3 Materials and Methods

3.3.1 Strains and growth conditions.

A table of relevant plasmids and strains is presented in Table B.1. All plasmids were cloned using Gibson assembly[56] or restriction digest cloning. Phosphomutants were constructed by mutating the phosphorylated residue to a glutamic acid to mimic a constitutively phosphorylated state or to an alanine to make it non-phosphorylatable. All phosphomutants were constructed using Q5 Site-Directed Mutagenesis kits following manufacturers instructions (New England Biolabs). *E. coli* Top10 strain (Life Technologies) was used in all experiments with two exceptions. For constructs with arabinose inducible promoters, the strain used was DP10, a previously described TOP10 variant with regulatable control of pBAD[88]. MG1655 *E. coli* strain was used in single cell experiments for ease of loading and growth in the microfluidic chamber. For KaiA overexpression, KaiA was driven by a pBAD promoter and KaiBC was driven by pTET (400ng/mL aTc induction). 500μM arabinose induction was used for overexpression of KaiA vs. 0μM induction for no overexpression. KaiC phosphorylation reporter strains contained the synthetic interaction dependent reporter, KaiC-α N-terminal domain (αNTD), and SasA-λCI. In order to avoid SasA being trapped in the phosphorylated state, RpaA and CikA, components involved in the de-phosphorylation of SasA were expressed for all timecourse experiments. Unless otherwise stated, cells were grown in LB media with appropriate antibiotics overnight, single colonies were picked into M9 minimal media with 0.4% glycerol, 0.1% casamino acids, antibiotics, and 500μM arabinose induction and/or 20μM IPTG induction, as necessary. All endpoint measurements were taken after overnight growth in inducing media.

3.3.2 Phos-tag SDS PAGE and western blot.

Cell lysates were prepared by incubation for 10 minutes at 100°C in 3% SDS. Prepared lysates were separated by SuperSep Phos-tag precast gels following manufacturers protocol (Wako Chemicals, Japan). After four washes, 10 minutes each, in transfer buffer (0.025M tris base, 0.192M glycine, 10% methanol) with 10mM EDTA and one additional wash in transfer buffer without EDTA, proteins were transferred to PVDF membranes using an iBlot Gel Transfer Device and iBlot Gel Transfer Stacks (Life Technologies, Grand Island, NY). Blotting was done using the SNAP-ID Protein Detection System (EMD Millipore, Bellerica, MA) following manufacturer's instructions with an additional overnight blocking step in 10% BSA to reduce background. Proteins of interest were probed with anti-FLAG (Abcam, ab1238) or anti-6xHis-tag (Abcam, ab1187) antibodies, both used at a final dilution of 1:10,000. Peroxidase conjugates were detected using SuperSignal West Dura Extended Duration substrate (Thermo Scientific). Quantification was performed in FIJI[144] using densitometry. Proportion of KaiC phosphorylation was determined by normalizing the intensity of the upper band (phosphorylated KaiC) to the sum of the intensities of both the upper and lower band (total KaiC).

3.3.3 Fourier analysis.

Discrete fourier transform was performed using a built-in MATLAB function. Unless otherwise stated, Butterworth bandpass filters were applied as previously described[96] to subtract background periodicities outside of the range of interest. The low pass filter was set at the measuring interval (1h) and the high pass filter set at the acquisition time of the experiment. When directly comparing the strength of the circadian periodicity at 24h between

control strains and clock strains, background subtraction of other periodicities outside of the circadian range (20-30h) was conducted by bandpass filtering.

3.3.4 Timecourse measurements.

Cells grown as described above were transferred to M9 minimal media with no carbon source for 1h to synchronize the population[16]. Synchronization of KaiC phosphorylation state occurs because the minimal media induces starvation, which dramatically lowers the ATP/ADP ratio. Since KaiC-P requires ATP, this effectively shifts most of the KaiC proteins into the unphosphorylated state. Following the minimal media shock, cells were then back diluted to OD600=0.1 into either 500mL (for KaiC phosphorylation analysis) or 200μL M9 minimal media (for plate reader analysis) with 0.5% succinate and 1mM leucine to promote slower growth and without induction to reduce population desynchronization. For analysis of KaiC phosphorylation, aliquots were taken at 4h intervals and lysed by incubation for 10 minutes at 100°C in 3% SDS and visualized on a western blot as described above. Mean-normalization was performed on each timecourse trace from the three biological replicates. Mean-normalized, average proportion of KaiC phosphorylation was plotted. For fluorescence timecourses, mineral oil was added to prevent evaporation. Readings were taken using a Synergy NEO HTS MultiMode microplate reader (Biotek, Winooski, VT) using appropriate filters. Fluorescence readings were normalized to OD600 and background (OD600 normalized fluorescence readings from a reporter-only strain AHC157) was subtracted. For both KaiC phosphorylation and synthetic reporter timecourses, statistical analyses were performed with RAIN[164] using a periodicity of 24.8h and over 3 days of the timecourse.

3.3.5 Imaging cells in a microfluidic device.

Cells were grown as described above to log phase (OD600 0.3), then resuspended into M9 minimal media with 0.5% succinate, 1mM leucine, and 0.85g/L pluronic 107, a surfactant to aid loading. The microfluidic chip was prepared, loaded, and imaged as previously described[123]. Image acquisition was performed once per hour using a custom MATLAB script and μManager[36] as described previously[123].

3.3.6 Single cell imaging data analysis.

Cell outlines were automatically determined using microbeTracker[155]. Average cell fluorescence of the mother cell (the bottom-most cell of each channel) was tracked over time. All subsequent analysis was performed using MATLAB. In order to examine frequencies of interest and make direct comparisons between the strength of the circadian periodicity in experimental and control samples, Butterworth bandpass filters were applied to traces as previously described[96], with the boundaries set as 20-30h, which are by definition circadian periodicities[95]. Discrete fourier transforms were performed and power was determined. The frequency corresponding to a 24h period was calculated to be 1.15×10^{-5} Hz. Power at that frequency was interpolated from the fourier spectra. A two-sample, one-sided KS-test was performed, comparing the population of cells with circadian clocks to control cells lacking the clock.

3.3.7 Endpoint measurements.

Cells were grown as described above. Fluorescence and OD600 were measured in triplicate on a Victor 3V 1420 Multilabel Counter (Perkin-Elmer) using appropriate filters. Fluores-

cence was normalized to OD. Across experiments, fluorescence was also normalized to the "no Δ" strain. Error bars are standard error of the mean (s.e.m.).

3.4 Results

KaiABC expressed in *E. coli* results in regulated phosphorylation of KaiC. To reconstruct the oscillator, we expressed FLAG-tagged KaiC in an operon with KaiA and KaiB in *E. coli* (Figure B.1A). Phos-tag^{TM} western blot was used to visualize KaiC-FLAG (Figure 3.1B-D). Two protein bands at 60kDa, corresponding to the size of KaiC monomers, were observed (Figure 3.1B). The upper band is phosphorylated KaiC (KaiC-P), as it does not appear in a non-phosphorylatable mutant (S431A, T432A) of KaiC[75] (Figure 3.1B). Moreover, we expressed *kaiA* and *kaiBC* under two distinct inducible promoters and observed that phosphorylated KaiC dominates (63%) in KaiA high-expression strains and that unphosphorylated KaiC dominates (42%) in strains with no KaiA expression (Figure 3.1C). Thus, overexpression of KaiA results in a higher proportion of phosphorylated KaiC (Figure 3.1C, $P<0.05$, Student t test) as would be expected from the known function of KaiA[76] (Figure 3.1A).

When KaiC phosphorylation was observed over time, it showed changes consistent with a circadian rhythm. We assayed KaiC phosphorylation (KaiC-P) by western blot in a strain expressing KaiABC and quantified the ratio of phosphorylated KaiC to total KaiC by densitometry. After ATP deprivation by transient minimal media shock (at t=0, see Methods) to synchronize KaiC-P levels, we observed oscillations in KaiC-P levels with a circadian period of approximately 24h for three days (Figure 3.1D, Figure B.2A). Phosphorylation levels over time were analyzed using RAIN[164] and circadian oscillations were shown to be sta-

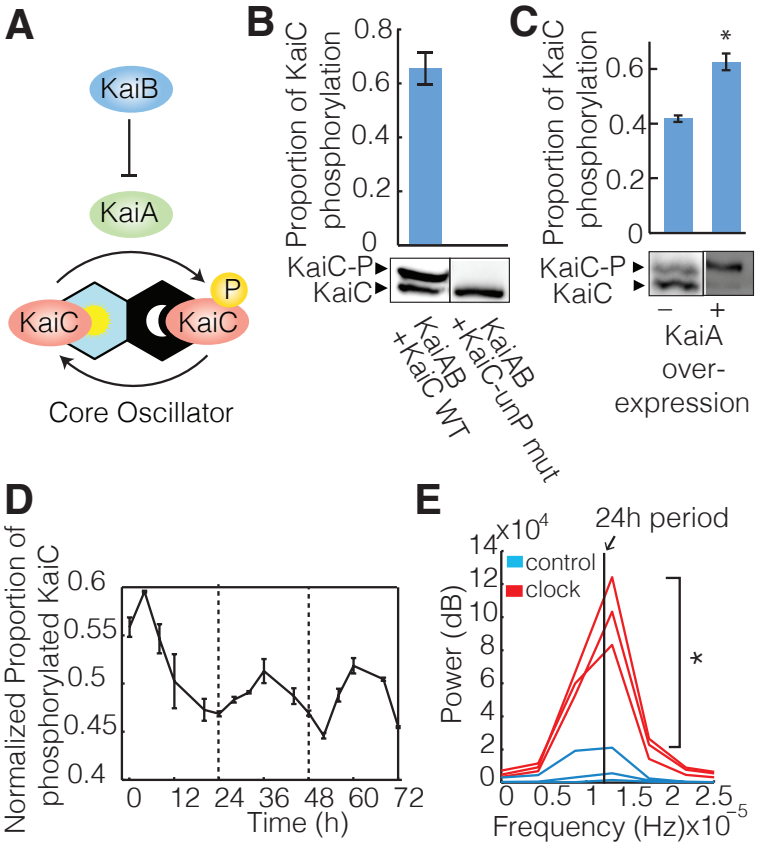

Figure 3.1: Functionality of the core KaiABC oscillator in a heterologous system. (A) Model for the phosphorylation state of KaiC through the antagonistic activities of KaiA and KaiB. (B) Levels of phosphorylated (-P) and unphosphorylated wild-type (WT) or non-phosphorylatable mutant (S431A, T432A, denoted "KaiC-unP mut") of KaiC-FLAG protein when expressed with KaiAB. Proteins were visualized by SDS-PAGE and immunoblotting (see Methods). The ratio of phosphorylated KaiC to total KaiC is plotted. (C) Levels of phosphorylated and unphosphorylated KaiC-FLAG, co-expressed with KaiB, with or without KaiA overexpression (see Methods). The ratio of phosphorylated KaiC to total KaiC is plotted. The asterisk indicates higher proportion of phosphorylated KaiC when KaiA is overexpressed ($P<0.05$, Student t test). (D) KaiC phosphorylation over time in *E. coli* co-expressing KaiA and KaiB, after synchronization (t=0h). The mean ratio of phosphorylated KaiC to total KaiC across biological replicates, mean normalized for each time-trace, is plotted. Circadian oscillations are statistically significant as analyzed using RAIN ($P<0.001$)[164]. (E) Power of circadian periodicity of KaiABC clock strains (red) and control strains expressing only KaiC (blue). Traces, produced from bandpass filtering for circadian periods (20-30h) followed by Fourier transform of time course data from biological replicates are plotted (n=3). Vertical line indicates the frequency corresponding to a 24h period. The asterisk indicates increased power of circadian periodicity in KaiABC clock strains ($p<0.05$, Student t test). Error bars, s.e.m. (n=3) in (B to D).

tistically significant (P<0.001). Proportion of unphosphorylated KaiC oscillates as well, but in antiphase as expected (Figure B.2B). Total KaiC expression showed no oscillating trends over time (RAIN, P=0.35) (Figure B.2C), and control strains expressing only KaiC without KaiA and KaiB showed increases in KaiC-P levels over time, but also no circadian trends (Figure B.3, RAIN P>0.99). Moreover, we compared the strength of the circadian periodicity at 24h between KaiABC clock strains and KaiC alone control strains (Figure 3.1E). To do so, first, background subtraction of other periodicities outside of the circadian range (20-30h)[95] was conducted by bandpass filtering. We then performed Fourier analysis, which allowed us to quantify the power of each trace at the circadian period (24h) and make direct comparisons between oscillation periods in kai clock and control strains. We found that the mean power at the 24h period of KaiABC clock strains was significantly higher than that of KaiC alone control strains (Figure 3.1E, Student t test p<0.05). These results indicate that KaiC phosphorylation oscillations, which required the presence of KaiA and KaiB, were circadian with periods of 24h.

In order to affect cellular physiology, a post-translational circadian clock needs to be connected to transcriptional output. To do so, we built a synthetic oscillator based on a modified bacterial two-hybrid system[29,30] (Figure 3.2A). It has been shown that phosphorylated KaiC binds more strongly than unphosphorylated KaiC to SasA, a downstream component to the core KaiABC oscillator[169] (Figure 3.2A inset). This property allowed us to build a synthetic transcriptional reporter of circadian state consisting of two binding partners: (a) full length KaiC C-terminally fused to the α subunit N-terminal domain of RNA polymerase (KaiC-αNTD), and (b) full length SasA C-terminally fused to λCI protein (SasA-CI) (Figure 3.2A, Figure B.4). Only if the two fusion proteins interact can they then bind

to the promoter and activate transcription of *gfp*.

Coexpression of the two binding partners, SasA-CI and KaiC-αNTD, with the remaining components of the core KaiABC oscillator, KaiA and KaiB, resulted in transcriptional outputs dependent on the presence of both binding partners as well as the phosphorylation state of KaiC. When compared to controls containing only KaiC-αNTD or only SasA-CI, gene expression from the synthetic reporter containing both fusion proteins was 3 fold higher after overnight induction (Figure 3.2B), indicating that both binding partners are necessary for reporter activation. A phosphomimic of KaiC showed higher reporter output compared to non-phosphorylatable KaiC, indicating that KaiC phosphorylation state determines reporter activity (Figure 3.2C). Interestingly, we also observed that changes in SasA phosphorylation could affect reporter output (Figure B.5), however, these changes were not necessary; KaiC phosphorylation changes alone were sufficient to change reporter output. These results (Figure 3.2B-C), taken together with our previous observation that KaiC phosphorylation state shows circadian oscillations (Figure 3.1D), suggests that KaiC circadian oscillations can be transmitted to a transcriptional output using this synthetic reporter.

Indeed, the transcriptional reporter showed circadian behavior over time; the synthetic circadian transcriptional system oscillated with a 24h period (RAIN, $P<0.05$, $n=3$) for more than three days after minimal media synchronization (Figure 3.2D). All biological replicates showed this trend, although there was some variability in phase and amplitude, likely due to lack of redundant mechanisms, such as those found in cyanobacteria[141,185], to maintain robustness and synchronization. Background reporter fluorescence alone showed linear increases over time without any circadian periods (Figure 3.2H). Fourier analysis

Figure 3.2 ***(following page)*****: A synthetic transcriptional reporter of KaiC phosphorylation demonstrates circadian oscillations.** (A) Schematic of the synthetic transcriptional reporter of KaiC phosphorylation state. Expression of the GFP reporter is dependent on the interaction between KaiC, which is fused to the α-subunit of RNA polymerase N-terminal-domain (KaiC-αNTD), and SasA, which is fused to λCI binding protein (SasA-CI). Inset: The binding of KaiC and SasA is dependent on phosphorylation state. Thickness of the arrow indicates strength of binding. (B) OD normalized fluorescent reporter output from cells expressing either binding partner alone or both KaiC-αNTD and SasA-CI. Error bars, s.e.m. (n=3) (C) OD normalized fluorescent reporter output of interactions between phosphomimic KaiC (S431A, T432E, -P mut), wildtype KaiC (WT), or non-phosphorylatable mutant of KaiC (S431A, T432A, -unP mut) with wild-type SasA-CI. Error bars, s.e.m. (n=3) (D to G) Background subtracted and OD normalized time course fluorescence when co-expressing KaiC (D) or KaiC S431A T432A non-phosphorylatable mutant (F), with SasA fusion proteins, and KaiAB. Brackets indicate 24h periods. Statistically significant circadian oscillations were observed in (D) using RAIN ($P<0.05$). Strength of the periodicity (power) at each frequency (and its corresponding period) is calculated using fourier analysis. Maximum power, shown by Fourier analysis of (D) and (F), occurs at a period of about 24 h in (E) and periods>30h and/or periods<20h in controls (G). Dashed line indicates a 24h period. Peaks align with dashed line in (D) and (F), but not in (G). t=0 indicates synchronization. (H) Background fluorescence, which was subtracted in D and F, from a strain containing only the reporter and no clock components.

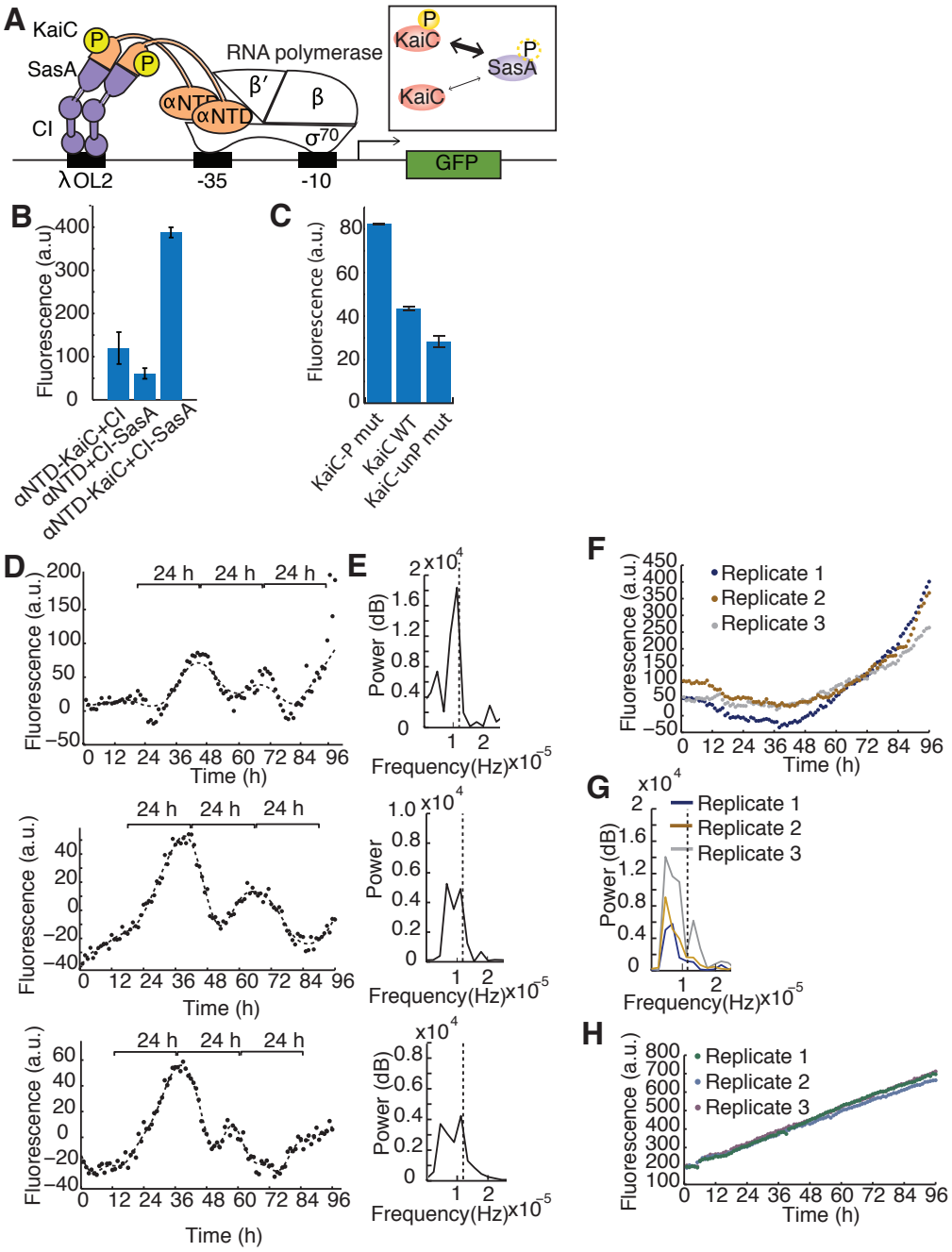

Figure 3.2: (continued)

revealed a dominant frequency with a period between 20-30h (Figure 3.2E). A negative control, consisting of a non-phosphorylatable mutant of KaiC with all other components remaining the same, showed fluctuations in fluorescence but no dominant period at 20-30h (Figure 3.2F-G).

Consistent with the population-wide circadian oscillations we observe, we also found that oscillations occur at the single cell level. Cells were loaded in channels of a microfluidic device that retains and analyzes only a single dividing 'mother' cell and not its progeny[171,123](Figure 3.3A-B, Video S4). This device is open on one end to the media flowing across the bottom, perpendicular to the channel (Figure 3.3A). Daughter cells were washed away, while mother cells were tracked across multiple cell divisions (Figure 3.3A). Using this method, single-cell fluorescence time trace data (Figure 3.3B-C) were collected for both circadian-circuit-containing cells and the corresponding control strains containing reporter alone. As above, to examine the frequencies of interest, a bandpass filter was applied to the data from both kai cock and control strains allowing for only signals from circadian periodicities, 20-30h[95] (Figure 3.3D). Fourier analysis provided a quantitative measurement, known as power, of the strength at each frequency, which allowed us to make direct comparisons between oscillation periods in kai clock and control strains. Kai clock strains showed an enrichment of high power signals with 24h periods compared to controls, which still show a small amount of power near this period likely due to cellular noise (Figure 3.3E). The overall distribution of power of the circadian clock cells was greater than that of controls (one-tailed, two sample Kolmogorov-Smirnov test, p=0.001) (Figure 3.3E), indicating that circadian oscillations at the population level are a result of single cell oscillations.

Having established circadian output of our synthetic device, we then implemented input

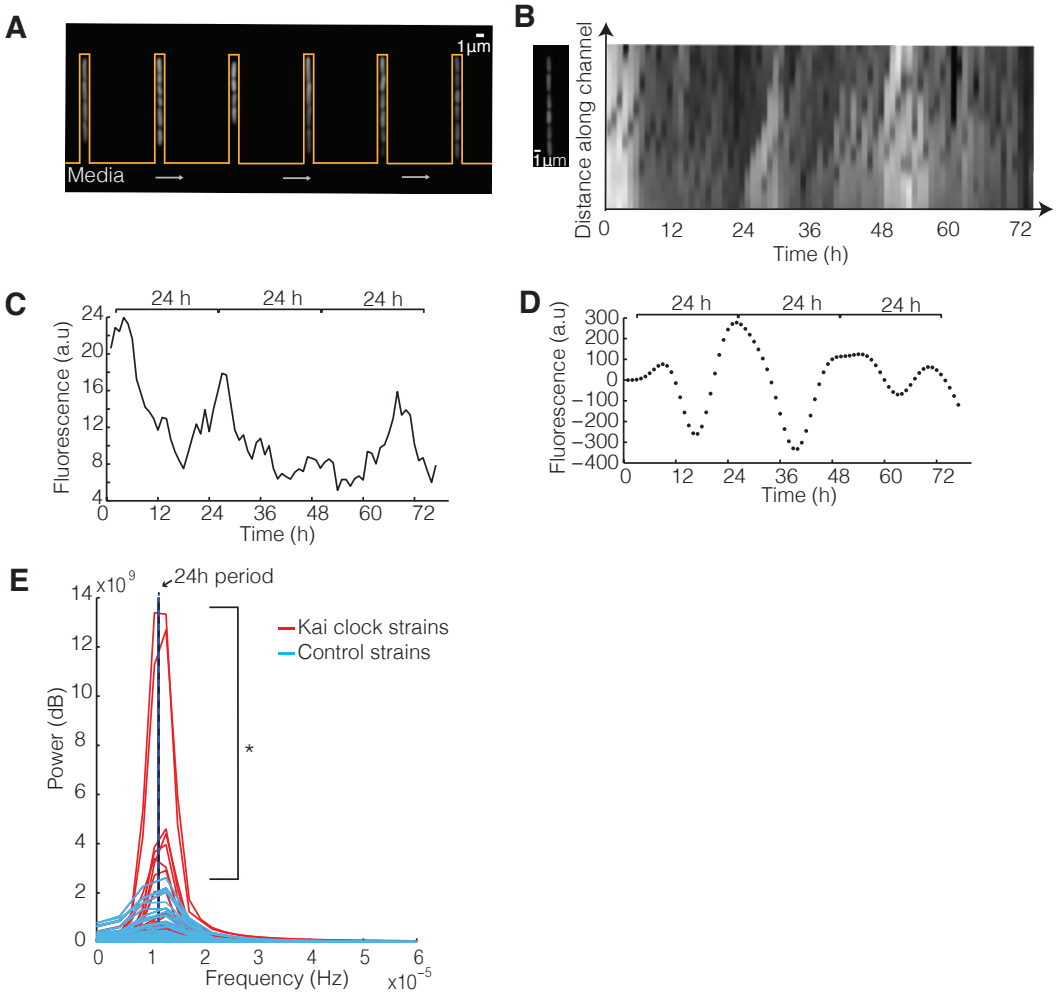

Figure 3.3: Single cells demonstrate oscillatory behavior. (A) Fluorescent cells in channels in the microfluidic device (orange outline) with media flow (grey arrows) across the base of the channels, allowing for long-term microscopy assay of single cell fluorescence. (B) Fluorescent cells in one channel of the microfluidic device at one time point (left). Kymograph shows fluorescence of a single channel over time (right), following overnight induction and minimal media shock synchronization (at t=0). Time interval: 1h. (C) Average fluorescence of a single mother cell, grown as in (A-B), containing the synthetic reporter and kai clock components. (D) Band pass-filtered (circadian periods included: 20-30h) data from (B) in order to compare the strength of circadian periodicity across multiple single cell traces. (E) Fourier spectra of the timecourses of kai clock strains (red, N=65) or control strains containing only the reporter (blue, N=48), both filtered as in (D). 11% of cells had a higher power than the control cell with the highest power. Asterisk indicates significantly different power of circadian periodicity in the population of cells containing the Kai clock compared to the control population ($P<0.05$, KS-test). Black line indicates the frequency corresponding to a 24h period.

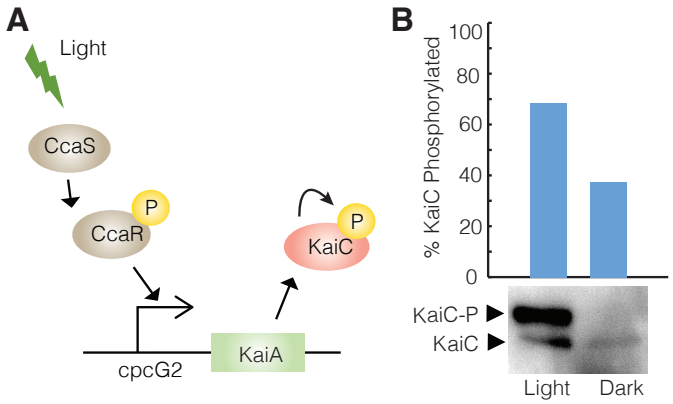

Figure 3.4: Light synchronization of the circadian clock. (A) Schematic of light input control of the kai clock (B) Percentage of phosphorylated KaiC after light induction or dark incubation assayed by western blot and quantified by densitometry. Size of bands: about 60kDa.

control of the circadian clock. To provide a robust and easy method of clock phase resetting and to mimic the natural mode of synchronization in cyanobacteria, light was chosen as the input control mechanism. KaiA was placed under the control of a light inducible system (Figure 3.4A)[160]. The light-inducible device, developed by Tabor *et al.*, utilizes a cyanobacterial two component system to activate a promoter upon green light absorption (Figure 3.4A). Light-inducible KaiA was co-expressed with KaiB and KaiC, and upon induction using light, 68.5% of KaiC was in the phosphorylated state (Figure 3.4B). When the same strain was grown in the dark, 37.4% of KaiC was in the phosphorylated state. This shows the potential to reset circadian phase through external light input.

After demonstrating a free-running oscillator with entrainability, we next decided to explore the outputs of the synthetic circadian clock. A post-translational oscillation signal can be transmitted to a natural output by introducing additional kai clock components and circadian responsive promoters from cyanobacteria into *E. coli* together with the basic KaiABC system. Several proteins including RpaA, SasA, and RpaB, have been implicated

in the transmission of the Kai signal (Figure 3.5A)[161,109,65,62]. We constructed *E. coli* strains (denoted "reporter+Kai clock") containing *kaiABC* in an operon driven by an arabinose-inducible promoter, a synthetic operon containing components *sasA, rpaA*, and *rpaB* driven by a lac promoter, and a fluorescent protein reporter driven by one of three circadian-responsive promoters[65] from *S. elongatus*: *kaiBC, rpoD6,* or *sigF2* (Figure 3.5A, Figure B.1). Induction overnight in unsynchronized cells resulted in reporter expression using each of the three promoters (Figure 3.5B). The kaiBC promoter (P_{kaiBC}) showed the highest fold-change (7.6-fold) between the reporter+kai clock strain and a strain with only the reporter. In contrast, the same reporter driven by the promoter of an *E. coli* housekeeping gene, ihfB, showed decreased expression when coexpressed with the kai clock and downstream components (Figure 3.5C), likely due to cellular burden when expressing heterologous components. These results indicate that the kai components are driving reporter expression from cyanobacterial promoters in the heterologous system.

Using this system, we identified a minimal cyanobacterial pathway capable of transducing KaiC phosphorylation activity to native *S. elongatus* circadian promoters in *E. coli*. We systematically deleted components from the reporter+kai clock strain with the *kaiBC* promoter driving reporter expression, (Figure 3.5D). When *kaiB* was deleted, reporter fluorescence decreased, indicating that KaiB was necessary for the increase in transcription, while a *kaiA* deletion resulted in no decrease in fluorescence (Figure 3.5D); thereby supporting the antagonistic activity of KaiA and KaiB on KaiC phosphorylation state[119] in this heterologous system. When *rpaB* was deleted, either alone or in combination with other components, reporter fluorescence was decreased. Previously, it was reported only that RpaB binds to circadian promoters[65]. Our data indicate that RpaB is necessary for transcriptional

Figure 3.5 *(following page)*: KaiC transmits phosphorylation information to cyanobacterial circadian promoters in *E. coli*. (A) Model pathway of cyanobacterial native circadian clock downstream components and their role, as identified by the current study. SasA binds to KaiC-P, is autophosphorylated, and phosphorylates the response regulator RpaA. RpaA antagonizes the activity of RpaB. RpaB then activates circadian-responsive-promoters. (B) Reporter fluorescence driven by cyanobacterial circadian-responsive promoters (*kaiBC*, *rpoD6*, or *sigF2*), with or without the kai clock components in (A). (C) Reporter fluorescence driven by an *E. coli* housekeeping gene promoter, with or without kai clock components. (D) Reporter fluorescence upon deletion of components from a strain expressing all components in (A). (E) Reporter fluorescence upon mutation to mimic phosphorylation (-P mut) or non-phosphorylatable (-unP mut) states. (F) Reporter fluorescence upon expression of CikA (see text). (G) OD normalized timecourse fluorescence of the mcherry reporter in a strain expressing all components shown in (A), after overnight induction and synchronization (t=0). In (D to G), reporter is P_{kaiBC} promoter driving mcherry expression and "no Δ" or "+kai clock" indicates fluorescence from a strain containing all components shown in (A). All measurements were taken after overnight induction. Error bars, s.e.m. (n=3) in (B to G).

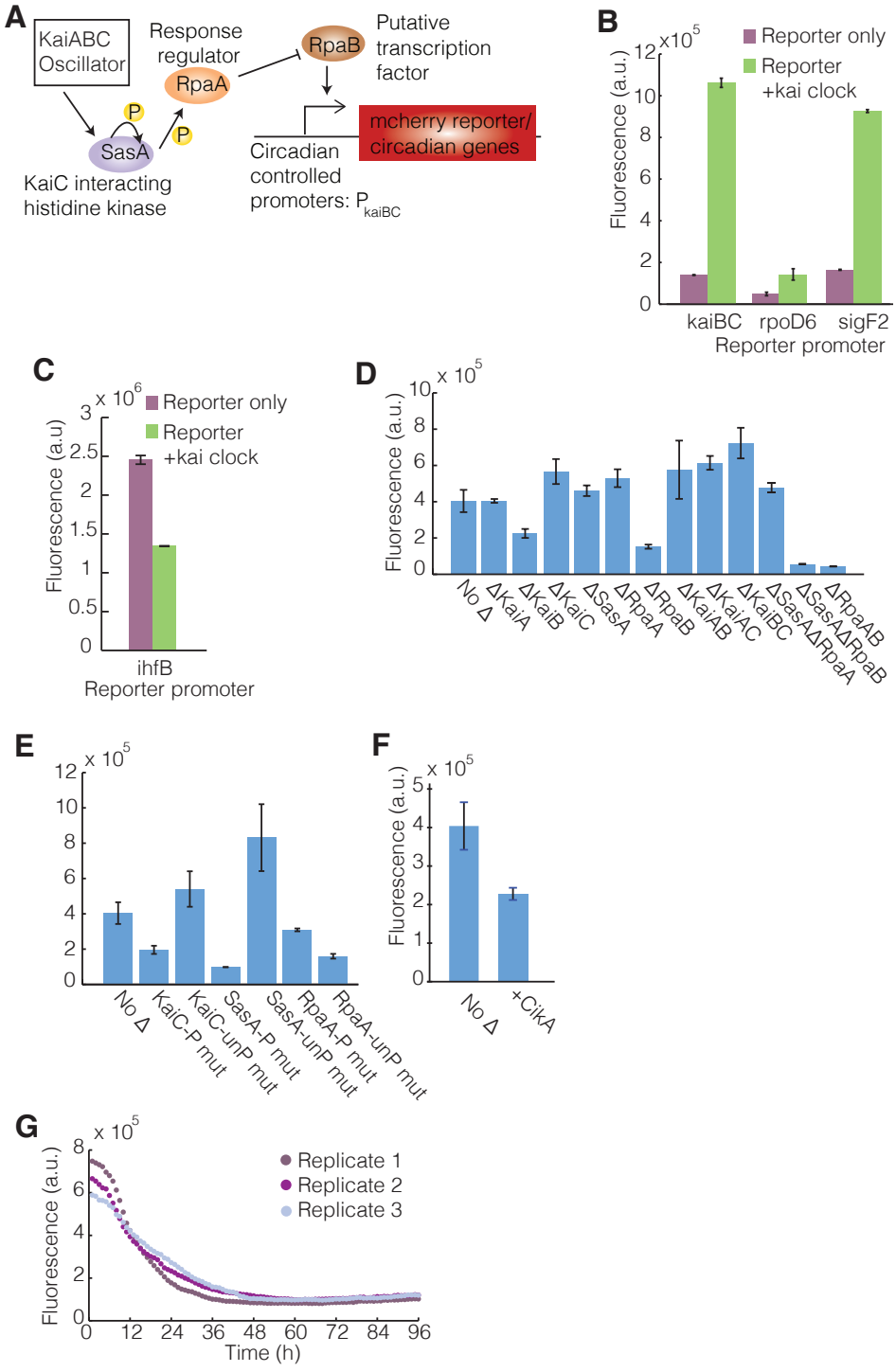

Figure 3.5: (continued)

activation by the core oscillator. This conclusion is corroborated by our other observations. When *sasA, rpaA* or *kaiC* were deleted, activation of the reporter increased or remained the same, consistent with a model where RpaB is an activator.

The phosphorylation states of KaiC and SasA affect reporter activation in our system (Figure 3.5E). A phosphomimic of KaiC or SasA decreased reporter fluorescence while a non-phosphorylatable mutation in KaiC or SasA resulted in increased reporter fluorescence. Surprisingly, RpaA phosphomimics showed a decrease in reporter fluorescence while non-phosphorylatable RpaA decreased reporter fluorescence further. These findings suggest that though the presence of RpaA is necessary for repression of RpaB, RpaA's phosphorylation state may not significantly affect its inhibition of RpaB. Additionally, inclusion of CikA, a phosphatase of RpaA[62], in the system also decreased reporter expression (Figure 3.5F). Taken together, these observations are consistent with RpaA also playing a direct role in activating the promoter[109] in our system. Alternatively, components upstream of RpaA may act directly on the promoter in an RpaA-independent manner. Furthermore, timecourses performed on the SasA/RpaA/RpaB strains with the P_{kaiBC}-driven reporter (Figure 3.5G) showed no oscillations. This suggests that additional components may affect the fidelity with which the kai clock transmits time information to an authentic cyanobacterial promoter as opposed to our synthetic system where we can reconstruct a minimal oscillator dependent on an *E. coli* rather than a cyanobacterial promoter. Models implicating global regulation of circadian gene expression by chromosome compaction in cyanobacteria[156] imply that necessary gyrases, sigma factors and other components are missing from *E. coli*, which has not evolved a similar global response.

3.5 Discussion

Our system demonstrates the transplantability of the bacterial circadian oscillator and its potential for re-engineering. While existing devices can program time delays[174] and produce cellular oscillators with relatively short periods[41], our oscillator has a period that matches the geophysical day-night cycle. We thereby improve the reach of temporal circuits in synthetic biology to encompass a biologically relevant periodicity. Recently, it was shown that disturbance of gut microbiota circadian rhythms, which are dictated by feeding rhythms, can result in host metabolic imbalances, such as obesity and glucose intolerance[165]. Modulation of gut microbes' circadian rhythms through engineering may therefore be a plausible strategy for correction of jet-lag induced dysbiosis. More broadly, our system will be pertinent as a platform for a variety of biotechnological applications, such as automated daily drug delivery, circadian control in industrial microbial processes, and treatment of circadian rhythm disorders[178,94].

4

Conclusion

The goal of this dissertation was to better understand cyanobacterial physiology for synthetic biology applications. With an interest in improving production of valuable compounds using cyanobacteria as a host, I attempted to understand their genome structure for better engineering and carbon fixation in order to increase photosynthetic efficiency. We would also like to take inspiration from cyanobacteria to engineer synthetic devices. To that end, I studied the assembly of a microcompartment capable of encapsulating reactions in a distinct microenvironment and I transplanted the circadian clock to a heterologous system. While the goals stated above are long term and not immediately achievable, the work presented in this dissertation represents progress towards these ends.

In chapter 1, I found that chromosomal duplication in cyanobacteria, a polyploid organ-

ism, is coupled to cell growth rather than cell division and that replication events occur in a spatially random fashion while segregation displays transient spatial ordering. These observations bring up several additional questions. How does the cell sense growth for deciding when to duplicate chromosomes? Does the duplication have checkpoints? One known checkpoint of cell division is the circadian rhythm[113], but what is the molecular mechanism that enforces that checkpoint? In addition, it would be interesting to understand what factors are responsible for spatial organization during the segregation process.

In order to truly solve the challenges surrounding cyanobacterial genome engineering posited in the introduction, one would need to turn this polyploid organism into a monoploid one. This would then enable selection to act against cells that don't acquire the gene of interest (with its corresponding antibiotic resistance marker) in a wide variety of engineering strategies. Rather than requiring constant and expensive antibiotic selection to maintain incompletely penetrant engineered chromosomes among a pool of wild-type chromosomes, monoploid cyanobacteri would require only round of selection to kill off any cells with non-engineered chromosomes. One potential strategy to make monoploid cyanobacteria would be to turn off duplication after it has occurred once. This would require a better understanding of the method of sensing cell growth and initiating duplication. Another more feasible solution that was made available after the publication of the work in Chapter 1 is to utilize the CRISPR/CAS system[134,23,106] to eliminate copies of the genome that didn't acquire the desired genes. All of these potential directions, built on the foundation of our work studying the polyploid genome organization in cyanobacteria, could lead to more efficient engineering of cyanobacteria for industrial purposes.

In Chapter 2, I found that a bacterial microcompartment, the carboxysome, is assembled

using the internal cargo, RuBisCO, as the primary organizer. This poses challenges for the assembly of synthetic cargo into these microcompartments. Previously, it was thought that a localization signal would be sufficient for targeting to the carboxysome. However, our work suggests that a simple signal is not responsible for localization to this microcompartment, perhaps explaining why the long search for a signal sequence has been unfruitful. Instead the biogenesis relies on fission, polyermerization and encapsulation of cargo.

Further work on the assembly process should focus on the specific self-assembly properties of RuBisCO (the carboxysome cargo) as well as the contacts between shell and cargo. Not only would this lead to a more complete understanding of the biogenesis process, but it could also lead to the ability to engineer the cargo and shell for increased carbon fixing capabilities. Thus far, engineering carbon fixation has focused on 1)engineering of the turnover rate and specificity of RuBisCO or 2) engineering heterologous carbon fixation pathways. I suggest that a third alternative might be engineering the carboxysome itself to encapsulate more cargo or adjust the permeability of the pores to increase the effect of the carbon concentrating mechanism. Mathematical modeling of the process suggests that optimizing these parameters may increase carbon fixation rates[107].

However, many challenges remain before engineering of the carboxysome will be feasible, especially if the goal is to encapsulate reactions other than carbon fixation. First, substrate entry and exit must be understood. In particular, it must be understood how the shell establishes or maintains selective permeability, or if such selectivity even exists. Also, the role of other cargo components, such as carbonic anhydrase, in the nucleation of shell assembly needs to be elucidated. The minimal set of components necessary to assemble the shell needs to be determined, and targeting of additional components to the inside of the

shell in a manner other than RuBisCO fusions needs to be found. After these challenges have been overcome, it might then become possible to engineer shells that can encapsulate heterologous reactions for the production of non-native compounds.

To expand on our results regarding the oxidizing internal microenvironment of the carboxysome, future studies should focus on understanding the other components that are unique to the carboxysome microenvironment, such as pH, and the mechanisms that establish the microenvironment. It has been suggested that optimal pH and redox potential is important for the functionality of the carbon fixation machinery. Further optimization may be done depending on the results of these studies. Finally, the mechanism underlying the creation of the microenvironment should be studied, in order to make it modifiable to suit the reaction conditions required of an engineered heterologous pathway.

In Chapter 3, I successfully engineered a synthetic device that produces a circadian rhythm in a non-circadian organism. These results not only further our understanding of the circadian rhythm in cyanobacteria, but also pave the way for the engineering of synthetic circadian rhythms for various medical applications. Further optimization can be performed in order to increase the robustness of the device, as the current system is highly susceptible to growth conditions and loss of synchrony. Future directions for this project should focus on tuning the inputs as well as connecting the device to useful outputs. I show preliminary data indicating function of a light controllable input to synchronize the synthetic circadian clock, however, further work needs to be done to tune the sensitivity of the input synchronization. In addition, the current output (GFP), while easy to visualize, is not particularly useful. Therefore, I recommend connecting other synthetic modules to the device. Counters, memory devices, and drug delivery vehicles are all examples of useful down-

stream modules that could be powerful when connected to a circadian clock for industrial biosensing and medical applications.

Synthetic biology has shown great promise for not only a better understanding of the complexities of biology, but also the design of faster and cheaper manufacturing processes, as well as better drugs and therapies. By studying native biological systems and using that knowledge to inform the design of synthetic biological devices, we can combine the power of these approaches to solve important problems that will transform manufacturing, healthcare, and the environment.

A

Supplemental Information for Chapter 2

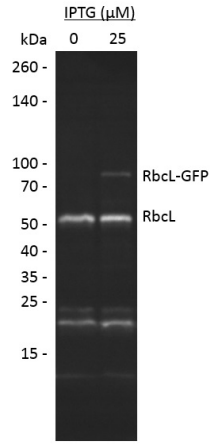

Figure A.1: Quantification of RbcL and RbcL-GFP levels by Western blot. The inducible RbcL-GFP strain was grown in the presence or absence of 25μM IPTG at early log phase for 12 hours. Using a rabbit polyclonal anti-RuBisCO antibody, the intensities of bands above background were quantified; the RbcL-GFP band is 11% of the intensity of the endogenous RbcL band.

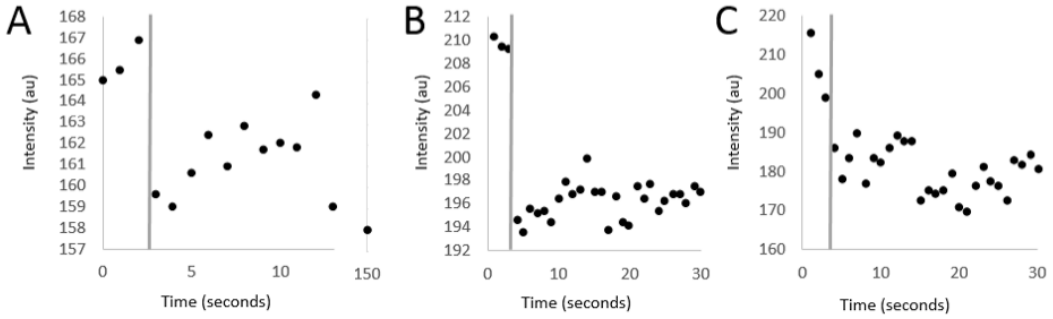

Figure A.2: Additional bar carboxysome FRAP data. Bleaching events are indicated by grey lines. Unbleached portions of the bar were used to correct for photobleaching.

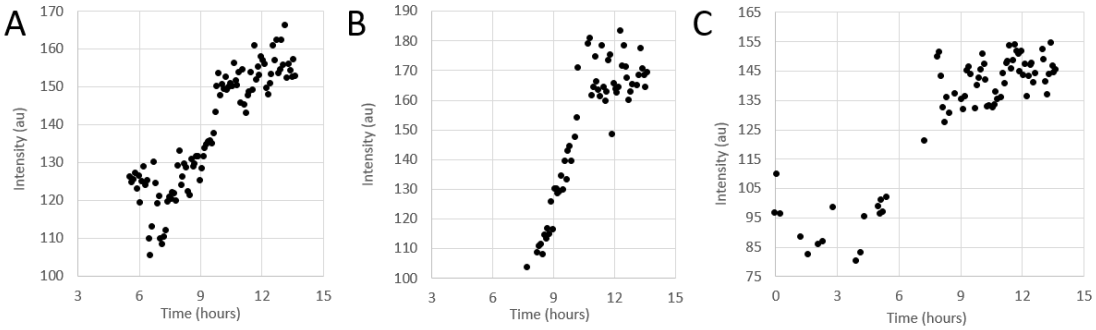

Figure A.3: Additional shell protein assembly data. (A–C) Individual traces of the fluorescence intensity of CcmK4 foci. Each panel represents a different cell, and only shell foci in the process of assembling are represented. Time interval: 5 minutes.

Table A.1: Strains and Plasmids

Strain or Plasmid	Relevant genotype	Resistance	Reference
E coli strains			
DH5-α	Host strain for plasmid construction		
S. elongatus strains			
PCC 7942	Wild-type *Synechococcus*, ATCC organism 33912		(Allen 1968)
RuBisCO/shell protein two color	Papca::ccmk4::sfGFP inserted in neutral site 1; lacI and ptrc::rbcL::mOrange2 inserted in neutral site 2	Kan/Sp	This work
Plasmids			
pDFS724	rbcL::sfGFP cloned into Neutral Site 2 at XbaI and NotI sites	Kan	(Savage 2010)
pDFS594S	Papca::ccmk4::YFP cloned into Neutral Site 1 at SpeI and NotI sites	Sp	(Savage 2010)
pAHC003	rbcL::mOrange2 cloned into pDFS724 in place of rbcL::sfGFP	Kan	This work
pAHC134	Papca::ccmk4::sfGFP	Sp	This work
pAHC126	rbcL::SNAP cloned into pDFS724 in place of rbcL::sfGFP	Kan	This work
pAHC149	rbcL::roGFP1 cloned into pDFS724 in place of rbcL::sfGFP	Kan	This work

Video S1.
Normal and bar carboxysomes are born from replicative events. Division events occur 2 seconds after the appearance of a white asterisk 2μm above the relevant carboxysome. Imaging was initiated 1 hour after induction. Green: RbcL-GFP. Red: phase contrast. Scale bar, 2μm. Frame rate, 12 frames (5 minute)/second.(AVI)

Video S2.
Shell protein is late to localize to RuBisCO assemblies. Shell localization events occur 2 seconds after the appearance of a white asterisk 1μm above the relevant carboxysome. The top-most highlighted carboxysome also nucleates a daughter at the 12 hour mark. Imaging was initiated 3 hours after induction. Green: RbcL-mOrange. Red: CcmK4-GFP. Scale bar, 2μm. Frame rate, 7 frames (5 minute)/second.(AVI)

Video S3.
Nascent, reduced carboxysomes oxidize as they mature. Oxidation events occur 2 seconds after the appearance of a white asterisk 1μm above the relevant carboxysome. Imaging was initiated 24 hours after induction. Green: 488Ex RbcL-roGFP1. Magenta: 410Ex RbcL-roGFP1. Scale bar, 2μm. Frame rate, 7 frames (10 minute)/second.(AVI)

B

Supplemental Information for Chapter 3

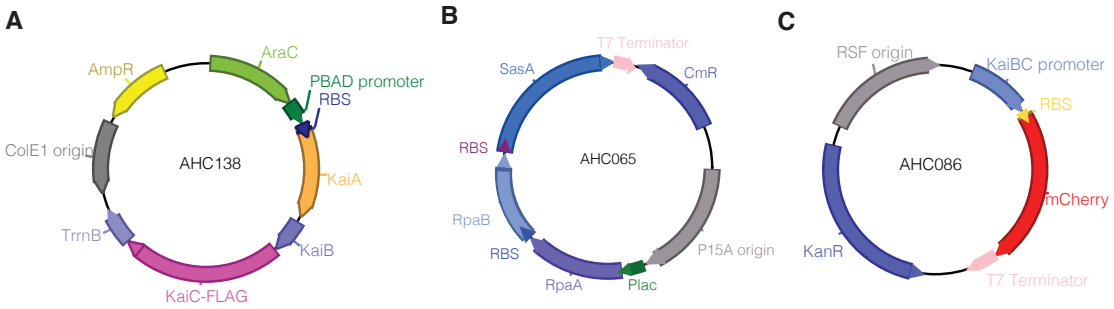

Figure B.1: Plasmids built for reconstruction of circadian oscillator in E. coli. (A) Core oscillator components, *kaiABC*, are expressed in an operon driven by an arabinose inducible promoter. (B) Additional native cyanobacterial components, *sasA*, *rpaA*, *rpaB*, are expressed in a synthetic operon driven by an IPTG inducible promoter. (C) *mCherry* reporter is driven by a circadian responsive promoter, such as *kaiBC*.

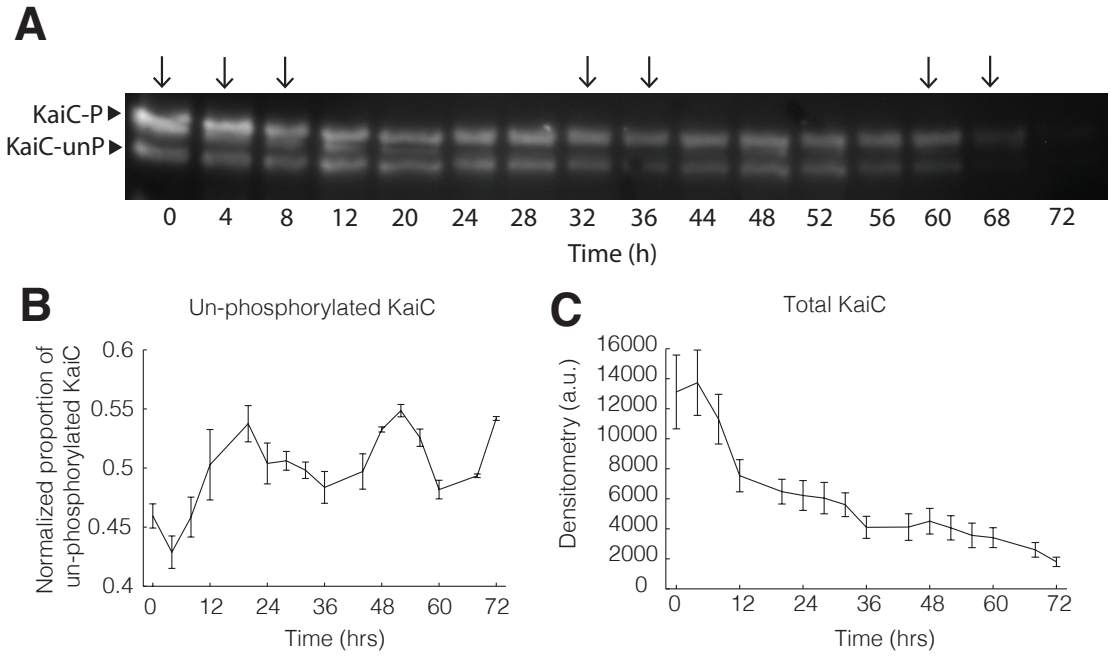

Figure B.2: Additional data and quantifications of KaiC phosphorylation in *E. coli* expression KaiABC. (A) Western blots, which were quantified in Fig. 1D, show phosphorylated and unphosphorylated KaiC over time in *E. coli* coexpressing KaiA and KaiB. Time t=0h corresponds to synchronization. Arrows indicate timepoints with high proportion of phosphorylated KaiC. (B) Proportion of unphosphorylated KaiC over time in *E. coli* co-expressing KaiA and KaiB, after synchronization (t=0h). The mean ratio of unphosphorylated KaiC to total KaiC across biological replicates, mean normalized for each time-trace, is plotted. (C) Total KaiC quantified over time in *E. coli* co-expressing KaiA and KaiB, after synchronization (t=0h). No statistically significant oscillations were found when analyzed using RAIN (P=0.35). Error bars, s.e.m. (n=3).

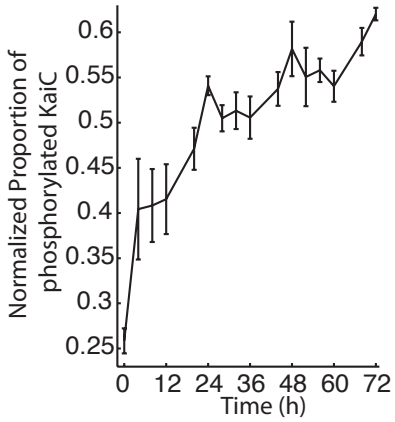

Figure B.3: Circadian phosphorylation of KaiC over time require KaiA and KaiB. KaiC phosphorylation over time, after synchronization, in *E. coli* expressing only KaiC without other Kai clock components. The mean ratio of phosphorylated KaiC to total KaiC across three biological replicates, mean normalized for each time-trace, is plotted. Error bars, s.e.m. (n=3). Circadian oscillations were not statistically significant as analyzed by RAIN (P>0.99).

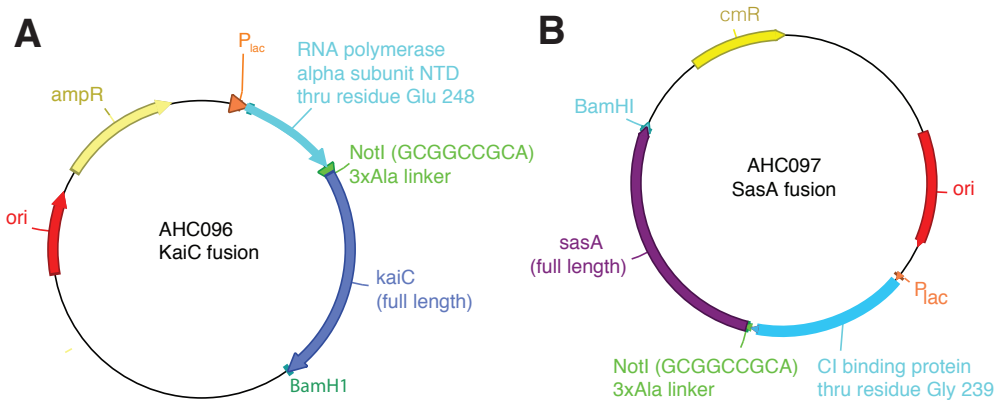

Figure B.4: Plasmids built for the synthetic oscillator based on a modified bacterial two-hybrid system. (A) Plasmid expressing full length KaiC C-terminally fused to the α subunit N-terminal domain of RNA polymerase (KaiC-αNTD). (B) Plasmid expressing full length SasA C-terminally fused to λCI protein (SasA-CI). The two parts of the fusions are connected via 3X alanine linkers and their corresponding genes are driven by lac promoters.

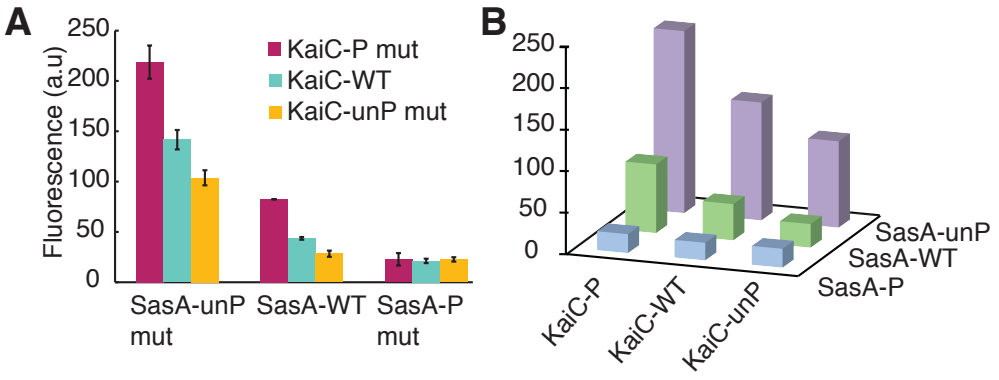

Figure B.5: KaiC and SasA phosphorylation states affect reporter output. (A) OD normalized fluorescent reporter output of interactions between combinations of phosphomimic (-P mut, KaiC: S431A, T432E; SasA: H161D), non-phosphorylatable mutant (-unP mut, KaiC: S431A, T432A; SasA: H161A) and wild-type (WT) KaiC-αNTD and SasA-CI. Error bars, s.e.m. (n=3) (B) 3D graph representation of the data presented in panel A.

Video S4.
Circadian rhythm seen in single cells in a microfluidic device.

Table B.1: Strains and Plasmids

Strain or plasmid	Relevant genotype	Resistance	Reference
E. coli strains			
DH10B/Top10	Expression strain		Invitrogen
MG1655	Parent strain of DH10B for microfluidic device loading and growth		
DP10	Variant of DH10B for arabinose induction		Kizer 2008
Plasmids			
AHC138	pBAD-kaiABC-FLAG	Amp	This work
AHC181	Phosphomutant variant of AHC138 with KaiC S431A, T432A	Amp	This work
AHC82	pTET kaiBC-6XHis	Kan	This work
AHC21	pBAD kaiA	Amp	This work
AHC20	pBAD-kaiABC	Amp	This work
AHC86	*kaiBC* promoter -mcherry	Kan	This work
AHC65	plac-rpaA, rpaB, sasA	Cm	This work
AHC22	AHC20 ΔkaiA	Amp	This work
AHC165	AHC20 ΔkaiB	Amp	This work
AHC166	AHC20 ΔkaiC	Amp	This work
AHC163	AHC65 ΔrpaA	Cm	This work
AHC164	AHC65 ΔsasA	Cm	This work
AHC123	AHC65 ΔrpaB	Cm	This work
AHC170	AHC20 ΔkaiAC	Amp	This work
AHC171	AHC65 ΔSasA ΔRpaB	Cm	This work
AHC172	AHC65 ΔRpaAB	Cm	This work
AHC85	AHC65 ΔSasA ΔRpaA	Cm	This work
AHC205	AHC20 KaiC phosphomimic S431D T432D		
AHC 140	AHC65 RpaA phosphomimic D53E	Cm	This work
AHC141	AHC65 Non-phosphorylatable RpaA D53A	Cm	This work
AHC 142	AHC65 RpaB phosphomimic D56E	Cm	This work
AHC 143	AHC65 Non-phosphorylatable RpaB D56A	Cm	This work
AHC 144	AHC65 SasA phosphomimic H161D	Cm	This work
AHC 145	AHC65 Non-phosphorylatable SasA H161A	Cm	This work
pBR α	P_{lacUV5}/P_{lpp} -directed synthesis of the full length a subunit of *E. coli* RNAP	Amp	Dove et al. 1997
pACλCI	P_{lacUV5}-directed synthesis of the lCI protein	Cm	Dove et al. 1997
AHC096	p*lacUV5*- and p*lpp*-directed synthesis of the αNTD (residues 1-248 of the α subunit of *E. coli* RNAP) fused via three alanines to *S. elongatus* KaiC protein	Amp	This work
AHC097	P_{lacUV5}-directed synthesis of the lCI protein fused via three alanines to *S. elongatus* SasA protein	Cm	This work
AHC157	p*lac*O_L2 driving 3x tandem sfGFP	Kan	This work
AHC177	AHC096 Non-phosphorylatable KaiC S431A, T432A	Amp	This work
AHC178	AHC096 KaiC phosphomimic S431D T432D	Amp	This work
AHC179	AHC097 Non-phosphorylatable SasA H161A	Cm	This work
AHC180	AHC097 Phosphomimic SasA H161D	Cm	This work

C

Chen, A. H., Silver, P. A. (2012). Designing biological compartmentalization. Trends in Cell Biology, 22(12), 662–70.

Review

Cell PRESS

Special Issue – Synthetic Cell Biology

Designing biological compartmentalization

Anna H. Chen[1] and Pamela A. Silver[1,2]

[1] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[2] Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA

Intracellular organization is a key factor in cell metabolism. Cells have evolved various organizational systems to solve the challenges of toxic pathway intermediates, competing metabolic reactions, and slow turnover rates. Inspired by nature, synthetic biologists have utilized proteins, nucleic acids, and lipids to construct synthetic organizational systems that mimic natural systems. Many of these systems have been applied to metabolic pathways and shown to significantly increase the production of industrially and commercially important chemicals. Further engineering and characterization of synthetic organizational systems will allow us to better understand native cellular strategies of spatial organization. Here, we discuss recent advances and ongoing efforts in designing and characterizing synthetic compartmentalization systems to mimic natural strategies and increase metabolic yields of engineered pathways.

Compartmentalization benefits natural and engineered systems

Biological complexity requires varying degrees of organization. Cells require spatial organization to perform the various enzymatic reactions and processes necessary to sustain life [1]. This is achieved through compartmentalization, the physical separation of biological reactions. Examples of compartmentalization include membrane-bound organelles, bacterial microcompartments [2,3], multienzyme complexes, and others [4,5].

Inspired by nature, synthetic biologists have recently devised strategies to mimic cellular organizational systems. These synthetic systems have been predominantly designed toward metabolic engineering of pathways, harnessing the capability of cells to produce industrially [6,7] or pharmaceutically useful [8,9] compounds.

In this review, we describe the various difficulties faced by the cell when performing metabolic reactions and natural compartmentalization systems that solve these problems. We review recent advances in designing synthetic compartments that provide modular solutions to overcome these same challenges. These systems capture the benefits of spatial organization and apply them to engineered pathways. This has also recently been reviewed in [10–12]. Our review will discuss the latest progress and challenges in designing compartmentalization, especially in building bacterial microcompartments and RNA scaffolds. We also analyze the degree to which these mimic natural systems and discuss how they aid in our understanding of the biological organization of the cell.

The need for intracellular organization

Cells face many challenges that benefit from compartmentalization (Figure 1a). First, some enzymes, such as ribulose 1,5-bisphosphate carboxylase oxygenase (RuBisCO) [13], suffer from slow turnover, which results in flux imbalances or bottlenecks in pathways. Reliance on such enzymes may require establishing local concentration gradients of substrates. This would increase reaction rates to support adequate pathway flux [14]. Second, diffusion of volatile intermediates through the cell membrane results in their loss from the cell [15]. Third, biosynthetic pathways can generate toxic intermediates that inhibit growth, such as hydrogen sulfide accumulated during bacterial sulfur metabolism [16]. Finally, metabolites can participate in multiple competing reactions, reducing their availability for any single pathway. An example of this is malonyl-CoA, an intermediate that is consumed in fatty acid and phospholipid production but is also used in the biosynthesis of polyketides and flavonoids [17].

Nature's solutions

To deal with these challenges, nature has evolved compartmentalization strategies (Figure 1b), such as large enzyme complexes [10,18,19] and organelles [2,20], to spatially organize metabolism. In eukaryotes, compartmentalization in the form of membrane-bound organelles is common. The peroxisome, for example, encapsulates reactions that generate or consume hydrogen peroxide, a toxic intermediate from the breakdown of organic substrates in oxidative reactions [21].

Until recently, prokaryotes were generally thought to lack internal organization [22]. However, researchers have recently discovered different types of bacterial microcompartments that partition the internal space of the bacterial cell for specialized functions [2,23]. In cyanobacteria and other autotrophic prokaryotes, carboxysomes encapsulate RuBisCO and carbonic anhydrase, enzymes involved in the rate-limiting step of the Calvin cycle [24,25] (Figure 2a). These proteinaceous microcompartments are the primary 'carbon-concentrating mechanism' in these bacteria. They

Corresponding author: Silver, P.A. (pamela_silver@hms.harvard.edu).
Keywords: compartmentalization; synthetic scaffolds; bacterial microcompartments; biological organization.

 0962-8924/$ – see front matter http://dx.doi.org/10.1016/j.tcb.2012.07.002 Trends in Cell Biology, December 2012, Vol. 22, No. 12

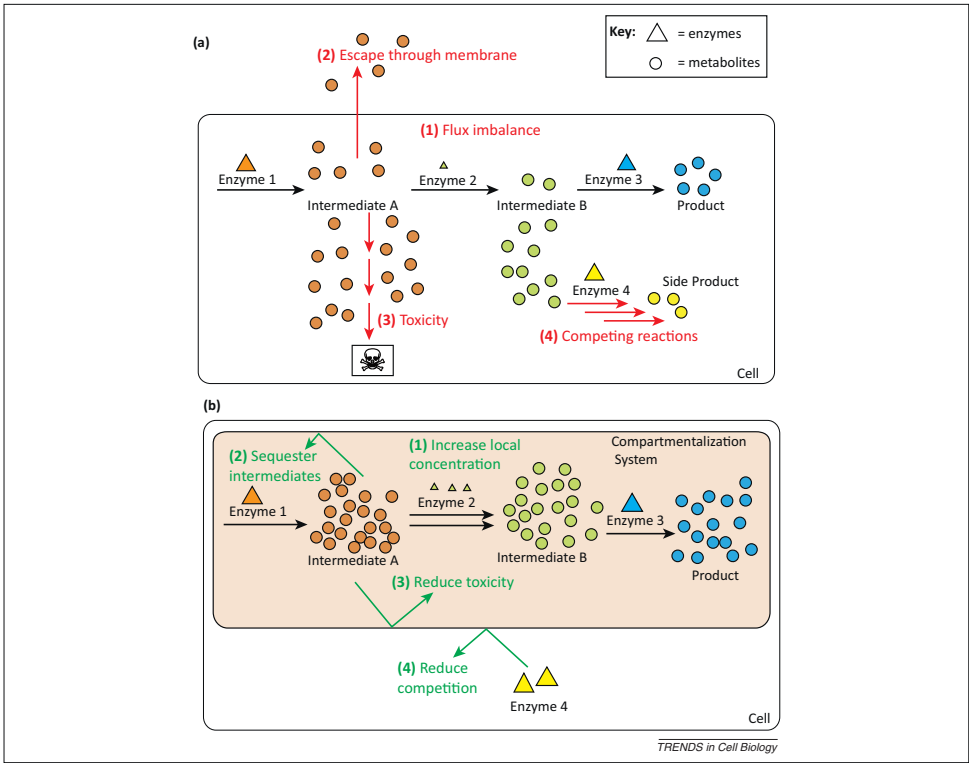

Figure 1. Compartmentalization: nature's solution to various challenges. **(a)** Nature faces many challenges when conducting the chemical reactions of the cell. (1) Differing enzyme kinetics may result in flux imbalances. (2) Intermediates may be lost through the cell membrane. (3) Toxic intermediates can result in growth inhibition. (4) Competing reactions can divert flux through undesired pathways. **(b)** Compartmentalization systems specifically solve challenges 1–4, respectively, by: (1) creating areas of local concentrations to favor reaction kinetics; (2) sequestering intermediates; (3) reducing toxicity; and (4) reducing competition. (Adapted from [11].)

are proposed to help overcome the slow turnover rate of RuBisCO by providing a high local concentration of carbon dioxide to the enzyme [26,27].

Two other bacterial proteinaceous microcompartments protect the cell from toxic aldehyde intermediates. The ethanolamine utilization (Eut) microcompartment sequesters acetaldehyde, a volatile and toxic intermediate of the ethanolamine utilization pathway [28]. Likewise, the 1,2-propanediol utilization (Pdu) microcompartment encapsulates propionaldehyde, minimizing its toxicity [29]. These and numerous other bacterial microcompartments have been found in approximately 400 microbial genomes [2].

Another method of compartmentalization found in nature is multienzyme complexes, which directly link enzymes involved in a given pathway. Ideally, this results in substrate channeling, the process by which intermediates are directly transferred between the active sites of two enzymes that catalyze sequential reactions in the pathway [30]. Substrate channeling prevents the loss of intermediates and minimizes competing cross-reactions. A classic example is tryptophan synthase, a multienzyme complex that catalyzes the last two reactions in the biosynthesis of L-tryptophan [18,31]. The intermediate, indole, is channeled from one active site to the next without being released into the surrounding environment. This is advantageous for the cell not only because indole is reactive and easily lost through the cell membrane, but also because, in the absence of indole, tryptophan synthase catalyzes dehydration of serine to pyruvate at 5% the rate of tryptophan formation [32]. Other multi-enzyme complexes found in nature include polyketide synthase [33], carbamoyl phosphate synthetase [34], and cellulosomes [35], which may function similarly by increasing reaction kinetics and reducing the loss of intermediates.

Synthetic compartmentalization

The goal of metabolic engineering is to optimize a given biosynthetic pathway to increase production of a particular substance [7,36]. Many of these pathways present the same challenges of toxic intermediates, competing reactions, and flux imbalances found in nature [10,14]. Therefore,

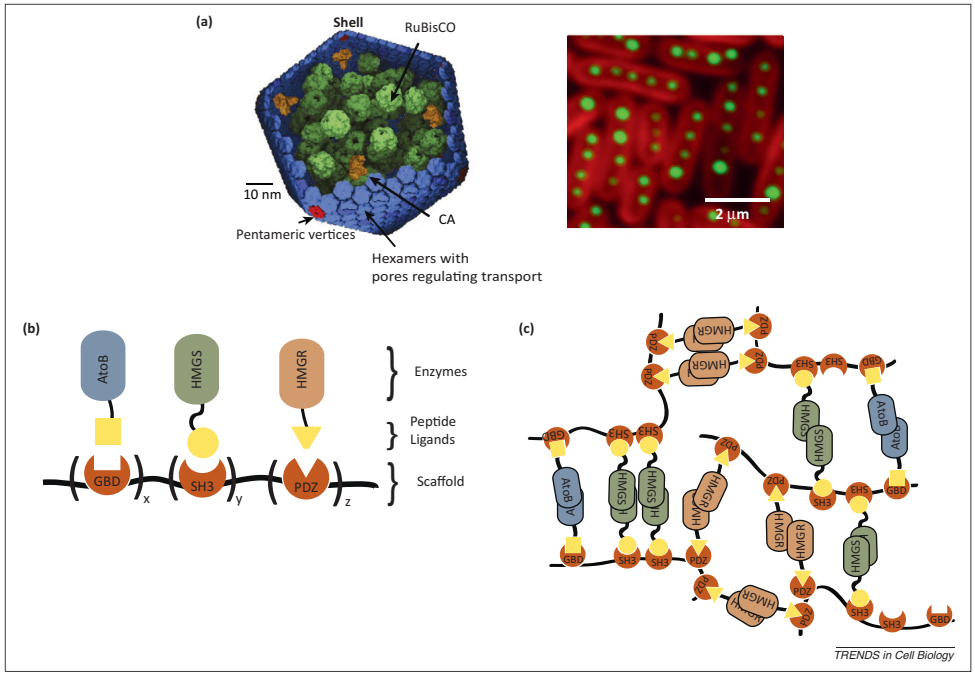

Figure 2. Protein scaffolds: utilizing nature's building material for functional complexes. Nature compartmentalizes reactions via enzyme complexes. Proteins often serve as scaffolds, bringing together pathway enzymes and increasing flux. **(a)** Bacterial microcompartments are protein structures that encapsulate reactions [2]. Carboxysomes contain ribulose 1,5-bisphosphate carboxylase oxygenase (RuBisCO) and carbonic anhydrase. The shell comprises pentamers and hexamers, some of which contain pores that are thought to regulate substrate transport (left). Carboxysomes spatially align in the cell [51] (right). Heterologous expression of microcompartments and further engineering of targeting and transport has the potential for making microfactories that may be capable of encapsulating various reactions. (Left: reprinted from [46] with permission from the authors. Right: image courtesy of David F. Savage, Berkeley, CA, USA). **(b)** Inspired by natural functional complexes, synthetic protein scaffolds were created [40]. Eukaryotic protein–protein interaction domains (GBD, SH3, PDZ) were expressed as a single polypeptide, creating a synthetic scaffold. Modular protein domain–ligand interactions were used to colocalize three enzymes of the mevalonate pathway: AtoB, HMGS, and HMGR. The stoichiometries of the enzymes could be controlled by changing the numbers (x, y, z) of each of the scaffold domains. Scaffolds balance metabolic flux, sequester a toxic intermediate, HMG-CoA, and create a high local concentration of intermediates, effectively mimicking nature's strategies of compartmentalization. **(c)** Scaffolds are hypothesized to oligomerize into large complexes that resemble metabolite microdomains, trapping intermediates in the interior and quickly converting them to product before they escape [11].

synthetic biologists have drawn inspiration from nature to design synthetic compartmentalization systems.

Compartmentalization by tethering using protein scaffolds

Nature uses proteins and protein–protein interactions to build functional multienzyme complexes. In some cases, non-catalytic scaffold proteins are used to assemble these complexes. One example of such a scaffold protein is Ste5, which selectively brings together MAPKK Ste7 and its substrate MAPK Fus3, promoting phosphorylation in this signaling cascade and disfavoring competing substrates [37]. Another example is scaffoldins, which organize protein subunits into the cellulosome and assist in the breakdown of cellulose [35]. Bacterial microcompartments, large assemblies comprising thousands of protein subunits, also function as organizational devices in the cell. Attempts to mimic natural protein scaffolding mechanisms have included crosslinking enzymes [38], enzyme immobilization on solid substrates *in vitro* [39], and direct enzyme fusion [10].

Inspired by naturally existing enzyme complexes, a synthetic protein scaffold, capable of colocalizing enzymes and increasing product titers, was produced [40] (Figure 2b). Eukaryotic protein–protein interaction domains were fused to form scaffolds. Each of the domains selectively recruited enzymes fused to cognate peptide ligands, colocalizing them to the scaffold.

Protein scaffolds have been tested on various pathways and increases in titers were shown. The mevalonate pathway, for example, produces a toxic intermediate and suffers from flux imbalance due to differing enzyme turnover rates. Scaffolding increased product yields by 77-fold [40]. The glucaric acid pathway was also scaffolded, and 5-fold improvement was seen over a high baseline titer of 0.6 g/L [41]. This pathway requires high levels of intermediate myoinositol to drive myoinositol oxygenase kinetics. A final example is biological hydrogen production, which suffers

from competing reactions and requires close contact of the two enzymes involved. This pathway benefited by about fivefold from protein scaffolding [42]. In principle, these synthetic protein scaffolds can replicate natural compartmentalization strategies found in multienzyme complexes, thereby alleviating the problems sometimes found in engineered metabolic reactions.

The extent of the increase in yields due to scaffolding depends on scaffold stoichiometry. Nature uses gene duplication events and evolution of protein–protein interactions, events that operate on a long timescale, to change the ratio of each of the enzymes in multienzyme complexes. Synthetic scaffolds similarly enable us to build and test many different enzyme stoichiometries, but in a short period of time, to determine the optimal ratio for production of a compound of interest. Varying the number of each interaction domain in the scaffold was shown to dramatically affect product titers [40]. In this way, designing stoichiometries and geometries of synthetic protein scaffolds may result in more effective compartmentalization mechanisms.

Although the exact mechanism by which protein scaffolds increase yields remains unknown, it has been hypothesized that they may form higher order complexes due to enzyme oligomerization [11] (Figure 2c), similar to naturally occurring metabolite microdomains, such as Ca^{2+} or cAMP microdomains [43,44]. If true, intermediates produced within the complex would be quickly consumed before they could escape into the cellular milieu, thereby reducing their toxicity. This would allow protein scaffolds to more closely approximate bacterial microcompartments.

Compartmentalization using protein encapsulation

Bacterial microcompartments are relatively closed to the surrounding environment with protein-based pores that likely regulate transport of substrate into and out of the proteinaceous shell (Figure 2a) [45]. This strategy effectively isolates reaction intermediates from the rest of the cellular milieu and could create high local concentrations of substrates, thereby increasing product yields.

Microcompartments have been heterologously expressed and could be used to encapsulate foreign pathways. As such, these microcompartments have the potential to increase pathway flux, making them useful tools for metabolic engineering. For example, carboxysomes from cyanobacteria (Figure 2a) were heterologously expressed in *Escherichia coli* and the encapsulated RuBisCO was functional in *in vitro* assays [46]. The Eut and Pdu microcompartments from *Salmonella* have also been expressed in *E. coli* [47,48].

Many challenges remain before heterologously expressed microcompartments can be utilized in metabolic engineering. Targeting of foreign enzymes to microcompartments is an area of ongoing research and localization sequences have been found for the Pdu and Eut microcompartments [47,49]. However, although associations between carboxysome shell proteins and other protein components have been shown [50], targeting novel enzymes to the carboxysome remains not well understood [46]. Another challenge is elucidating the mechanism of substrate transport across the microcompartment shell. In carboxysomes, diffusion of small molecules across the shell is likely, but pores that suggest active regulation of substrate crossing have also been revealed in crystal structures [45]. Heterologously expressed microcompartments, isolated from their native cellular environment, may help us better understand these mechanisms.

Spatial organization of microcompartments can be regulated within a cell, providing the cell with another mechanism to optimize metabolic processes. For instance, carboxysomes are aligned along the major axis of the cell [51] (Figure 2a, right) and the bacterial cytoskeletal component ParA is required for proper alignment to occur. The exact interaction between the cytoskeleton and the bacterial microcompartment remains unknown and is an active area of research. This knowledge would enable the design and construction of a novel synthetic cytoskeletal scaffolding device. By inducing polymerization and depolymerization of cytoskeletal proteins in response to substrate availability or other environmental signals, we may be able to control the spatial organization and number of compartments tethered to this synthetic cytoskeleton. This may allow dynamic regulation of metabolic reactions in synthetic microcompartments.

Another bacterial microcompartment-like encapsulation device was recently engineered. A lumazine synthase capsid from *Aquifex aeolicus* was heterologously expressed in *E. coli* [52]. Lumazine synthase, an enzyme catalyzing riboflavin synthesis, forms icosahedral assemblies that do not naturally encapsulate other enzymes. However, electrostatic interactions were engineered, enabling it to encapsulate a toxic enzyme, HIV protease. Directed evolution was used to select for a capsid with higher loading capacity. In addition, the assembly state of lumazine synthase was controlled through engineered point mutations, which changed the quaternary structure of the capsid, thereby changing the size and structure of the capsid, without disrupting the secondary or tertiary structure [53]. The ability to insulate toxicity from the cytoplasm using a designable compartmentalization system is important for metabolic engineering. It allows the expression of heterologous pathways that may otherwise be toxic, or even lethal, to the cell.

Overall, protein assemblies have been used effectively to organize metabolic reactions and increase product yields. Thus far, however, such increases have been low. Therefore, the present challenge is to further increase yields to levels high enough for industrial applications. To do so, the mechanism of scaffold action must be further elucidated. In addition, scaffolds have been applied to only a few select pathways as proof of principle. Generality and scalability of scaffolds will require a better understanding of localization to synthetic scaffolds.

Compartmentalization using nucleic acids

DNA and RNA nanotechnology is a field of research that has many promising applications in medicine and industry [54–56]. Short strands of DNA or RNA can fold into various structures or assemble into dimer or multimer building blocks (Figure 3c). These building blocks can then polymerize into various 3D structures *in vitro*, including simple structures such as sheets, as well as more complex

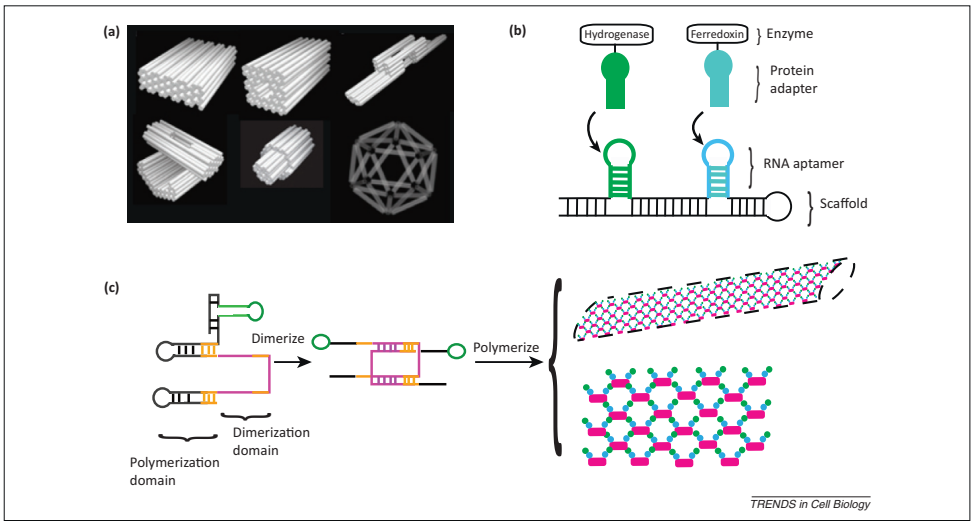

Figure 3. Nucleic acid scaffolds: mimicking the diverse geometries found in nature. (**a**) DNA and RNA nanotechnology can produce many different structures *in vitro,* including tubes and sheets [86]. Although nucleic acids are not frequently used in nature for scaffolding, the ease with which synthetic biologists can design and build these structures make them a versatile tool for building synthetic scaffolds that can mimic various natural compartmentalization strategies of different geometries. (Adapted, with permission, from [86].) (**b**) Synthetic RNA scaffolds were designed *in silico* and expressed *in vivo* [75]. Binding between RNA aptamers and protein adapters facilitated recruitment of enzymes to the scaffold. Hydrogenase and ferredoxin, components of a hydrogen production pathway, were scaffolded and pathway output was increased. (**c**) Single strands of RNA with aptamers dimerize and these building blocks polymerize into 2D sheets and nanotubes. Polymerization domains are protected with hairpins before dimerization to prevent self-binding and assure order of assembly. The versatile architectures of RNA nanostructures *in vivo* allow the mimicking of various natural compartmentalization systems.

structures such as tubes and capsules [57,58] (Figure 3a). In this way, DNA and RNA nanotechnology can be used to build complexes that structurally mimic natural organizational systems, such as enzyme channels and bacterial microcompartments.

Nucleic acid structures can also functionally mimic natural organizational systems. That is, proteins can be targeted directly to the scaffolds. Several multienzyme systems have been successfully organized using nucleic acids *in vitro*. NADH-flavin mononucleotide (FMN) oxidoreductase and luciferase, enzymes that catalyze two consecutive reaction steps, were assembled onto DNA scaffolds using streptavidin–biotin linkages and coupled enzymatic activity was increased relative to unscaffolded enzymes [59]. Glucose oxidase (GOX) and horseradish peroxidase (HRP) were immobilized through covalent conjugation to DNA oligonucleotides and an increase in reactivity was seen in both scaffolded systems compared with unscaffolded enzymes alone [60]. Other enzyme cascades have been organized using large supramolecular DNA scaffolds [61] and DNA origami [62] and reactivity was found to depend on relative enzyme positioning on the scaffold.

Although traditional *in vitro* assembly of DNA or RNA nanostructures relies on physiologically unsustainable temperatures and slow cool down to denature and anneal designed strands correctly [63], isothermal strategies for both DNA and RNA were recently developed [64–66]. This paved the way for *in vivo* applications of RNA and DNA scaffolds.

A DNA scaffold, in the form of a plasmid, was constructed. It recruited enzymes via zinc finger domains that bound to specific motifs on the scaffold [67]. Several enzymes were tested using this system, including pathways for the production of resveratrol, 1,2-propanediol, and mevalonate. Yields were found to increase as a function of scaffold architecture, similar to protein scaffolds.

Scaffolds with more complex geometries were constructed using designed non-coding (nc)RNA (Figure 3b,c). As a building material, RNA has several advantages. Because the base-pairing interactions in RNA (and DNA) are well studied [68,69], secondary and tertiary structures of nucleic acids can be predicted *in silico* [70–72] with nanometer precision [57] and are relatively easy to design [55]. RNA can also be expressed in large amounts by engineered cells and can stably exist outside the nucleus in eukaryotic cells, making it portable to a wide range of hosts. Additionally, RNA can recruit proteins fused to adapters via RNA aptamers, which can be evolved to bind with high specificity to various ligands [73,74].

A single molecule of RNA (approximately 100 bases) can fold into a linear discrete scaffold that mimics multienzyme complexes in nature [75] (Figure 3b). The RNA molecule comprises multiple different aptamer motifs, which can recruit proteins, as well as complementation regions, which form the scaffold (Figure 3b). This scaffold is relatively easy to build and characterize and is suitable for expression of specific stoichiometries of enzymes in a chain [75]. Scale up, which may allow production at industrially

feasible titers, is also possible; numerous RNA aptamer domains can be connected in a single molecule of RNA. In this way, almost 100 proteins were recruited in a chain for visualization experiments [76], although this has not yet been adapted for metabolic engineering purposes.

An RNA nanotube was built using short single-stranded building blocks that polymerized *in vivo* [75] (Figure 3c). Here, the directionality of enzyme loading was not specified; that is, enzymes were tethered to both the interior and the exterior of the nanotube. However, loading enzymes selectively to the interior has been shown *in vitro* [77]. If such selective loading could be accomplished *in vivo*, a modified version of this 1D scaffold could act like a pipe, preventing intermediates from escaping and increasing product yields. Another scaffold architecture, a 2D sheet, was constructed using a similar strategy and product yield was highest when using this scaffold (Figure 3c). Although there is no directly analogous system in nature, this 2D-scaffold is likely to function by increasing the number of neighbors any enzyme has in close proximity, allowing exploration of the quantitative effect of colocalizing additional proteins in two dimensions; for example, by recruiting them to a membrane.

As a demonstration of feasibility, [FeFe]-hydrogenase and ferredoxin, enzymes involved in the biological production of hydrogen [78], were recruited to RNA scaffolds. This pathway is a useful test bed because it suffers from competing reactions and requires the enzymes to be in close proximity [42]. Hydrogen output was found to increase as a function of scaffold architecture.

DNA and RNA scaffolds allow metabolic engineers to mimic many different strategies implemented by nature for substrate channeling and reducing competition to increase pathway flux. *In vitro* work provides a wide array of tested architectures and functionalities that mimic various natural compartmentalization systems in different ways. This ability to design precise scaffold geometry is a major advantage of nucleic acid scaffolds over protein and lipid compartmentalization systems, where structure prediction and design remain difficult. However, the *in vitro* to *in vivo* transition is a major challenge for nucleic acid scaffolds and is not always possible due to many constraints, such as isothermal assembly. Specific control over scaffold architectures and the ability to design customized shapes with RNA and DNA isothermally *in vivo* are crucial next steps in utilizing this technology to increase yields. If successful, RNA and DNA may even be used to build scaffolds in architectures not found in nature. Although the effects of such structures on the cell's metabolic load and the crosstalk with native systems may be problematic, it raises interesting possibilities for metabolic engineers to exceed nature's capabilities to produce useful compounds.

RNA scaffolds currently have an added advantage over DNA and other types of scaffold in that the amount of scaffold expression can be controlled. However, the effect of expressing high amounts of nucleic acid on cell growth and viability is not well understood and titration of scaffold expression has yet to be performed.

The generality of nucleic acid scaffolds remains to be demonstrated, because they have been used in only a few select pathways. Further application of nucleic acid scaffolds will require generation of additional orthogonal RNA aptamers or zinc finger domains. The application of RNA scaffolds, with its diverse architectures, to various pathways will not only allow further improvement of the technology, but will also allow testing of the mechanism of scaffold action. For example, the effect of scaffolding is hypothesized to be greater for non-diffusible intermediates compared with diffusible intermediates if local high concentration of intermediates is an important contributing factor. Overall, RNA scaffolds are capable of becoming versatile, designable scaffolds that can mimic natural systems of differing geometries, compartmentalize various metabolic pathways, and increase yields, possibly beyond nature's capabilities.

Compartmentalization using lipids

Lipids, often in the form of membranes, are widely used to encapsulate reactions in nature. Lipid vesicles and oil emulsions have been used to perform a wide variety of reactions *in vitro*, such as gene expression [79], sequencing, and evolution of new enzymes [80]. Even self-replicating lipid systems have been created: protocells that are capable of catalyzing RNA reactions and processing 'food' micelles [81]. Although these systems are effective at encapsulation of reactions, targeting specific enzymes to synthetic lipid vesicles remains difficult because localization signals or other specific targeting mechanisms to lipid vesicles are unknown.

Another compartmentalization strategy utilizes naturally existing membrane-bound organelles. Unlike lipid vesicles, many membrane-bound organelles have well studied localization mechanisms [82]. One study targeted methyl halide transferase to yeast vacuoles to increase methyl iodide production [83]. The increases seen were likely to have been due to added access to a cofactor and sequestration of halogenated products. Also, targeting of terpenoid production pathway components to the mitochondria significantly increased yields [84]. The advantage of this system is that it allows engineering of eukaryotic cells such as commercially important yeast.

Lipid compartmentalization systems remain a relatively unexplored area of synthetic biology, because they have many weaknesses compared with nucleic acid or protein scaffolds. For example, controlling the architecture of compartments and stoichiometries of encapsulated enzymes is challenging. Other major challenges for *in vivo* use of lipid compartmentalization devices include secretion of final products and, in the case of native organelle utilization, potential interactions with the host metabolism. A better understanding of these factors and studies in natural lipid compartmentalization systems may yield results that can inform the development of synthetic lipid compartments.

Using synthetic scaffolds to understand the limitations of natural systems

Studying engineered scaffolds may give us new insight into the function and design principles of natural compartmentalization strategies. For example, scaffold architecture and enzyme stoichiometries were found to be important

to pathway output [40,75]. Through scaffolding pathways with known stoichiometries and differing architectures, it may be possible to probe the limitations of native systems and even improve on them.

Knocking out a native compartment, replacing it with a synthetic one, and studying the resulting effects on the cell may allow us to decouple biological functions. This will help elucidate the functions and limitations of the native system. Conversely, studying heterologous expression of natural compartments and targeting will allow us to determine the components that are necessary and sufficient for proper function. Mechanisms of localization, assembly, and substrate gating may be elucidated from such studies.

In addition, the well-studied examples of natural compartmentalization systems have comprised primarily proteins and lipid membranes. Synthetic systems have shown, for example, that nucleic acids can function as scaffolds. However, RNA and DNA are not commonly used as materials for scaffolding in nature. Yeast telomerase is one of a few known examples of RNA serving as a flexible scaffold for protein subunits [85]. Given findings of the capabilities of a synthetic RNA scaffold, it may be that an as yet to be discovered function of ncRNA, which constitutes 80% of total RNA, is scaffolding. Synthetic scaffolds, in this case, may help us elucidate the organizational function of the material from which the scaffold was made.

Concluding remarks

Cells have evolved strategies for regulating and maintaining proper metabolic flux. Synthetic biologists have taken inspiration from nature and developed synthetic systems that solve some of the same problems. From direct mimicry using proteins to more novel solutions using nucleic acids, synthetic biologists have devised promising strategies for increasing metabolic production of industrially useful chemicals. However, some major challenges remain (Box 1). For example, scaffolds must be applicable to various pathways outside the test pathways chosen in the proof-of-principle studies conducted so far. Further increases in yields must be obtained before scale up to industrially relevant quantities can be achieved. Utilizing multiple types of scaffold in the same system, if feasible, may help accomplish this, because improvements may be made orthogonally by each scaffold. In addition, the effect of such high levels of production on cell growth must be minimized. To overcome these challenges, a better understanding of scaffold mechanism of action and enzyme localization must be achieved. If these efforts prove successful, scaffolds can be widely applied as an orthogonal novel method of increasing yields in metabolic engineering efforts. In addition, implementing and studying synthetic systems will help us understand the native biological organization of the cell.

Acknowledgments

The authors thank D. Ducat, C.J. Delebecque, J.E. Dueber, G. Sachdeva, D.F. Savage, J. Torella, and G.C. Wu for discussions and helpful suggestions. This work is supported by the National Institutes of Health (Grant #GM080177) and a Graduate Research Fellowship from the National Science Foundation to A.H.C. and by the Wyss Institute for Biologically Inspired Engineering and Department of Defense to P.A.S.

Box 1. Outstanding questions

- What is the mechanism by which scaffolds operate to increase yield? Do they form higher order structures?
- How can we predictably engineer scaffold architecture and control assembly?
- For what pathways are different scaffolds most useful? Is there a pathway size limitation? Does the diffusability of intermediate substrates influence the effectiveness of scaffold action?
- How do substrates enter and exit synthetic compartments? How are enzymes localized?
- How can we further improve yields using synthetic scaffold technologies? How do we scale up the existing technologies to industrial scale production?
- Are there other undiscovered natural scaffolding systems that utilize proteins, nucleic acids, and lipids? How does biological organization function natively?

References

1 Hurtley, S. (2009) Spatial cell biology. Location, location, location. Introduction. *Science* 326, 1205
2 Kerfeld, C.A. *et al.* (2010) Bacterial microcompartments. *Annu. Rev. Microbiol.* 64, 391–408
3 Kerfeld, C.A. *et al.* (2005) Protein structures forming the shell of primitive bacterial organelles. *Science* 309, 936–938
4 Berlatzky, I.A. *et al.* (2008) Spatial organization of a replicating bacterial chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14136–14140
5 Komeili, A. (2012) Molecular mechanisms of compartmentalization and biomineralization in magnetotactic bacteria. *FEMS Microbiol. Rev.* 36, 232–255
6 Savage, D.F. *et al.* (2008) Defossiling fuel: how synthetic biology can transform biofuel production. *ACS Chem. Biol.* 3, 13–16
7 Lee, S.K. *et al.* (2008) Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* 19, 556–563
8 Chang, M.C.Y. and Keasling, J.D. (2006) Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* 2, 674–681
9 Khalil, A.S. and Collins, J.J. (2010) Synthetic biology: applications come of age. *Nat. Rev. Genet.* 11, 367–379
10 Conrado, R.J. *et al.* (2008) Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr. Opin. Biotechnol.* 19, 492–499
11 Lee, H. *et al.* (2011) Spatial organization of enzymes for metabolic engineering. *Metab. Eng.* 14, 242–251
12 Agapakis, C.M. *et al.* (2012) Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat. Chem. Biol.* 8, 527–535
13 Eichelmann, H. *et al.* (2009) Rubisco in planta kcat is regulated in balance with photosynthetic electron transport. *J. Exp. Bot.* 60, 4077–4088
14 Jørgensen, K. *et al.* (2005) Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant Biol.* 8, 280–291
15 Manjasetty, B.A. *et al.* (2003) Crystal structure of a bifunctional aldolase-dehydrogenase: sequestering a reactive and volatile intermediate. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6992–6997
16 Liu, Y. *et al.* (2012) Methanogens: a window into ancient sulfur metabolism. *Trends Microbiol.* 20, 251–258
17 Xu, P. *et al.* (2011) Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metab. Eng.* 13, 578–587
18 Hyde, C.C. *et al.* (1988) Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem.* 263, 17857–17871
19 Pfeifer, B.A. and Khosla, C. (2001) Biosynthesis of polyketides in heterologous hosts. *Microbiol. Mol. Biol. Rev.* 65, 106–118
20 Martin, W. (2010) Evolutionary origins of metabolic compartmentalization in eukaryotes. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 365, 847–855
21 Gehrmann, W. and Elsner, M. (2011) A specific fluorescence probe for hydrogen peroxide detection in peroxisomes. *Free Radic. Res.* 45, 501–506

22 Losick, R. and Shapiro, L. (1999) Changing views on the nature of the bacterial cell: from biochemistry to cytology. *J. Bacteriol.* 181, 4143–4145
23 Yeates, T.O. *et al.* (2008) Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat. Rev. Microbiol.* 6, 681–691
24 Tanaka, S. *et al.* (2008) Atomic-level models of the bacterial carboxysome shell. *Science* 319, 1083–1086
25 Cannon, G.C. and Shively, J.M. (1983) Characterization of a homogenous preparation of carboxysomes from *Thiobacillus neapolitanus*. *Arch. Microbiol.* 134, 52–59
26 Tanaka, S. *et al.* (2009) Insights from multiple structures of the shell proteins from the beta-carboxysome. *Protein Sci.* 18, 108–120
27 Badger, M.R. (2003) CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* 54, 609–622
28 Penrod, J.T. and Roth, J.R. (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*. *J. Bacteriol.* 188, 2865–2874
29 Sampson, E.M. and Bobik, T.A. (2008) Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *J. Bacteriol.* 190, 2966–2971
30 Zhang, Y-HP. (2011) Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* 29, 715–725
31 Dunn, M.F. (2012) Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Arch. Biochem. Biophys.* 519, 154–166
32 Xu, Y. and Abeles, R.H. (1993) Inhibition of tryptophan synthase by (1-fluorovinyl)glycine. *Biochemistry* 32, 806–811
33 Tran, L. *et al.* (2010) Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. *Chem. Biol.* 17, 705–716
34 Thoden, J.B. *et al.* (1997) Structure of carbamoyl phosphate synthetase: a journey of 96 A from substrate to product. *Biochemistry* 36, 6305–6316
35 Fontes, C.M.G.A. and Gilbert, H.J. (2010) Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu. Rev. Biochem.* 79, 655–681
36 Keasling, J.D. (2010) Manufacturing molecules through metabolic engineering. *Science* 330, 1355–1358
37 Zeke, A. *et al.* (2009) Scaffolds: interaction platforms for cellular signalling circuits. *Trends Cell Biol.* 19, 364–374
38 Sheldon, R. (2007) Enzyme immobilization: the quest for optimum performance. *Adv. Synth. Catal.* 349, 1289–1307
39 Mosbach, K. and Mattiasson, B. (1970) Matrix-bound enzymes. II. Studies on a matrix-bound two-enzyme-system. *Acta Chem. Scand.* 24, 2093–2100
40 Dueber, J.E. *et al.* (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759
41 Moon, T.S. *et al.* (2010) Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metab. Eng.* 12, 298–305
42 Agapakis, C.M. *et al.* (2010) Insulation of a synthetic hydrogen metabolism circuit in bacteria. *J. Biol. Eng.* 4, 3
43 Laude, A.J. and Simpson, A.W.M. (2009) Compartmentalized signalling: Ca2+ compartments, microdomains and the many facets of Ca2+ signalling. *FEBS J.* 276, 1800–1816
44 Baillie, G.S. (2009) Compartmentalized signalling: spatial regulation of cAMP by the action of compartmentalized phosphodiesterases. *FEBS J.* 276, 1790–1799
45 Klein, M.G. *et al.* (2009) Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *J. Mol. Biol.* 392, 319–333
46 Bonacci, W. *et al.* (2012) Modularity of a carbon-fixing protein organelle. *Proc. Natl. Acad. Sci. U.S.A.* 109, 478–483
47 Choudhary, S. *et al.* (2012) Engineered protein nano-compartments for targeted enzyme localization. *PLoS ONE* 7, e33342
48 Parsons, J.B. *et al.* (2010) Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Mol. Cell* 38, 305–315
49 Fan, C. *et al.* (2010) Short N-terminal sequences package proteins into bacterial microcompartments. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7509–7514
50 Kinney, J.N. *et al.* (2012) Elucidating the essential role of the conserved carboxysomal protein CcmN reveals a common feature of bacterial microcompartment assembly. *J. Biol. Chem.* 287, 17729–17736
51 Savage, D.F. *et al.* (2010) Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science* 327, 1258–1261
52 Wörsdörfer, B. *et al.* (2011) Directed evolution of a protein container. *Science* 331, 589–592
53 Chen, H-N. and Woycechowsky, K.J. (2012) Conversion of a dodecahedral protein capsid into pentamers via minimal point mutations. *Biochemistry* 51, 4704–4712
54 Pinheiro, A.V. *et al.* (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.* 6, 763–772
55 Guo, P. (2010) The emerging field of RNA nanotechnology. *Nat. Nanotechnol.* 5, 833–842
56 Famulok, M. and Ackermann, D. (2010) RNA nanotechnology: inspired by DNA. *Nat. Nanotechnol.* 5, 634–635
57 Aldaye, F.A. *et al.* (2008) Assembling materials with DNA as the guide. *Science* 321, 1795–1799
58 Wei, B. *et al.* (2012) Complex shapes self-assembled from single-stranded DNA tiles. *Nature* 485, 623–626
59 Niemeyer, C.M. *et al.* (2002) DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H:FMN oxidoreductase and luciferase. *Chembiochem* 3, 242–245
60 Müller, J. and Niemeyer, C.M. (2008) DNA-directed assembly of artificial multienzyme complexes. *Biochem. Biophys. Res. Commun.* 377, 62–67
61 Wilner, O.I. *et al.* (2009) Enzyme cascades activated on topologically programmed DNA scaffolds. *Nat. Nanotechnol.* 4, 249–254
62 Numajiri, K. *et al.* (2010) Discrete and active enzyme nanoarrays on DNA origami scaffolds purified by affinity tag separation. *J. Am. Chem. Soc.* 132, 9937–9939
63 Castro, C.E. *et al.* (2011) A primer to scaffolded DNA origami. *Nat. Methods* 8, 221–229
64 Afonin, K.A. *et al.* (2010) In vitro assembly of cubic RNA-based scaffolds designed in silico. *Nat. Nanotechnol.* 5, 676–682
65 Jungmann, R. *et al.* (2008) Isothermal assembly of DNA origami structures using denaturing agents. *J. Am. Chem. Soc.* 130, 10062–10063
66 Yin, P. *et al.* (2008) Programming biomolecular self-assembly pathways. *Nature* 451, 318–322
67 Conrado, R.J. *et al.* (2012) DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acids Res.* 40, 1879–1889
68 Jaeger, L. and Chworos, A. (2006) The architectonics of programmable RNA and DNA nanostructures. *Curr. Opin. Struct. Biol.* 16, 531–543
69 Rothemund, P.W.K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* 440, 297–302
70 Zadeh, J.N. *et al.* (2011) NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* 32, 170–173
71 Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415
72 Andronescu, M. (2003) RNAsoft: a suite of RNA secondary structure prediction and design software tools. *Nucleic Acids Res.* 31, 3416–3422
73 Ellington, A.D. and Szostak, J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346, 818–822
74 Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249, 505–510
75 Delebecque, C.J. *et al.* (2011) Organization of intracellular reactions with rationally designed RNA assemblies. *Science* 333, 470–474
76 Golding, I. and Cox, E.C. (2004) RNA dynamics in live *Escherichia coli* cells. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11310–11315
77 Douglas, S.M. *et al.* (2007) DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6644–6648
78 Ducat, D.C. *et al.* (2011) Rewiring hydrogenase-dependent redox circuits in cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3941–3946

79 Nomura, S.M. *et al.* (2003) Gene expression within cell-sized lipid vesicles. *Chembiochem* 4, 1172–1175
80 Channon, K. *et al.* (2008) Synthetic biology through biomolecular design and engineering. *Curr. Opin. Struct. Biol.* 18, 491–498
81 Zhu, T.F. and Szostak, J.W. (2009) Coupled growth and division of model protocell membranes. *J. Am. Chem. Soc.* 131, 5705–5713
82 Siddiqui, M.S. *et al.* (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS Yeast Res.* 12, 144–170
83 Bayer, T.S. *et al.* (2009) Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* 131, 6508–6515
84 Farhi, M. *et al.* (2011) Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab. Eng.* 13, 474–481
85 Zappulla, D.C. and Cech, T.R. (2004) Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10024–10029
86 Douglas, S.M. *et al.* (2009) Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* 459, 414–418

D

Chen, A. H., Afonso, B., Silver, P. A., Savage, D. F. (2012). Spatial and temporal organization of chromosome duplication and segregation in the cyanobacterium Synechococcus elongatus PCC 7942. PloS One, 7(10), e47837.

OPEN ACCESS Freely available online

PLOS ONE

Spatial and Temporal Organization of Chromosome Duplication and Segregation in the Cyanobacterium *Synechococcus elongatus PCC 7942*

Anna H. Chen[1☯], Bruno Afonso[1☯], Pamela A. Silver[1,2*], David F. Savage[3]

1 Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts, United States of America, **3** Department of Molecular and Cell Biology and Department of Chemistry, University of California, Berkeley, California, United States of America

Abstract

The spatial and temporal control of chromosome duplication and segregation is crucial for proper cell division. While this process is well studied in eukaryotic and some prokaryotic organisms, relatively little is known about it in prokaryotic polyploids such as *Synechococcus elongatus PCC 7942,* which is known to possess one to eight copies of its single chromosome. Using a fluorescent repressor-operator system, *S. elongatus* chromosomes and chromosome replication forks were tagged and visualized. We found that chromosomal duplication is asynchronous and that the total number of chromosomes is correlated with cell length. Thus, replication is independent of cell cycle and coupled to cell growth. Replication events occur in a spatially random fashion. However, once assembled, replisomes move in a constrained manner. On the other hand, we found that segregation displays a striking spatial organization in some cells. Chromosomes transiently align along the major axis of the cell and timing of alignment was correlated to cell division. This mechanism likely contributes to the non-random segregation of chromosome copies to daughter cells.

Citation: Chen AH, Afonso B, Silver PA, Savage DF (2012) Spatial and Temporal Organization of Chromosome Duplication and Segregation in the Cyanobacterium *Synechococcus elongatus PCC 7942.* PLoS ONE 7(10): e47837. doi:10.1371/journal.pone.0047837

Editor: Mark Isalan, Center for Genomic Regulation, Spain

Received August 12, 2012; **Accepted** September 21, 2012; **Published** October 24, 2012

Funding: This work is supported by the National Institutes of Health (Grant #GM080177) and a Graduate Research Fellowship from the National Science Foundation to AHC, by the Fundação para a Ciência e a Tecnologia and Graduate Program in Areas of Basic and Applied Biology (GABBA) to BA and by the Wyss Institute for Biologically Inspired Engineering and Department of Defense to PAS, and the Department of Energy Office of Science Early Career Research Program through Office of Basic Energy Sciences Grant DE-SC0006394 to DFS. DFS was a Department of Energy Physical Biosciences Fellow of the Life Sciences Research Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pamela_silver@hms.harvard.edu

☯ These authors contributed equally to this work.

Introduction

Genomic DNA replication and segregation are fundamental processes crucial to survival for all organisms. This process has been well studied in many bacterial species, including *Escherichia coli* [1], *Bacillus subtilis* [2–4], and *Caulobacter crescentus* [5,6]. Most of these organisms possess a single copy of one, two or three different chromosomes (Fig. 1A, I-III). In contrast, the cyanobacterium *Synechococcus elongatus PCC 7942* has multiple copies of its single chromosome – estimates suggest between three to six copies [7,8] (Fig. 1A, IV). To date, little is known about the dynamics of replication and segregation in prokaryotes with multiple copies of a single chromosome.

Most studies of replication and segregation so far have been conducted in monoploid bacterial species. In many of these organisms, replication timing and synchrony is strictly regulated [9]. In *E. coli*, for example, all origins fire synchronously at a fixed cell size per origin (initiation mass) that is independent of the growth rate [5,10]. Synchrony is tightly coupled to cell division cycles and ensures that daughter cells receive the correct number of chromosomes. However, regulating timing of replication may not be as important for proper cell division in polyploid organisms.

In addition to timing of replication, the spatial localization of replication is important for proper cell division. In *E. coli*, newly synthesized chromosomes appear in the cell center or at the quartile points along the long axis of the cell [11]. Replication forks appear at the origin of replication, separate into two sister replisomes that migrate to opposite cell halves as replication proceeds and returns to mid-cell as replication ends [11]. In *C. crescentus*, replisomes move towards the middle from the cell poles [12]. In this study, we probe the localization dynamics of chromosome replication in *S. elongatus*. We also investigate the stringency of the spatial organization of replication.

It has been suggested that organisms with multiple chromosomes do not require an active segregation mechanism, since given a large number of chromosomes, it is likely that each daughter cell will receive at least one copy [4]. This is analogous to high-copy plasmid systems, which typically lack an active segregation mechanism [13].

In order to better understand chromosome replication and segregation in the polyploid organism, *S. elongatus*, we tagged and visualized the chromosome and the replisome using a fluorescent repressor-operator system. Chromosome count and localization data was collected. We probed the spatial organization of *S. elongatus* chromosome segregation and found that, contrary to

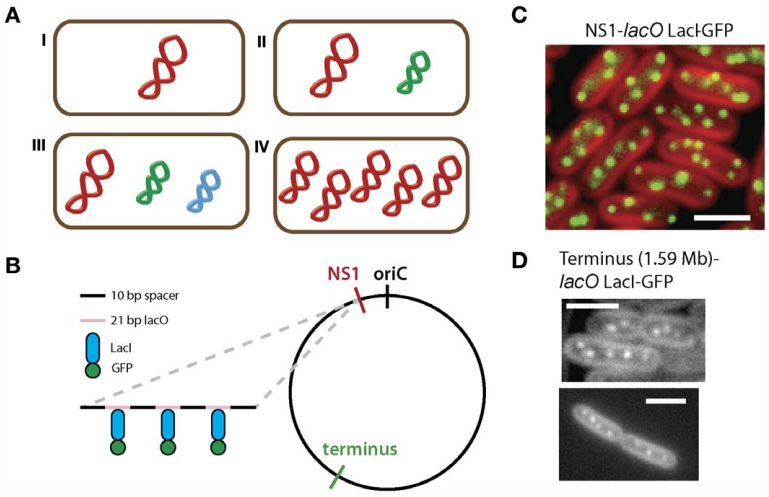

Figure 1. Chromosomes in the polyploid bacterium *S. elongatus* can be visualized using a fluorescent repressor-operator system. (A) Bacteria contain different genomic arrangements. Here, each color represents a different chromosome. They can possess a single copy of one chromosome (I), or have multipartite genomes (II-III) with one large chromosome (red) and one or more smaller chromosomes (green and blue). Some species of bacteria, such as cyanobacteria *Synechococcus elongatus* PCC 7942, are polyploid. That is, they have multiple complete copies of one chromosome (IV). (B) Chromosomes can be tagged and observed *in vivo* using a fluorescent repressor-operator system. *lacO* arrays were integrated either near the origin of replication (NS1) or the predicted terminus in the *S. elongatus* chromosome. 10 bp spacers with random sequences were inserted between the operator sites to avoid recombination (black). The protein fusion LacI-GFP (blue and green) bound to multiple repeats of its cognate *lacO* operator site (pink). (C) The fluorescent repressor-operator system from (B) was transformed into *S. elongatus*. The origins of replication of each chromosome appear as foci (green) in cells (red) when imaged using wide field fluorescence microscopy. Origins of replication are seen throughout the cell. (D) Cells with *lacO* arrays integrated near the putative terminus region at 1.59 Mb in the genome were visualized. Foci appear throughout the cell, similar to (C).
doi:10.1371/journal.pone.0047837.g001

previously suggested models, a surprising alignment occurs during the process. In addition, we calculated the timing of replisomes and the diffusive dynamics of replication. We found that chromosome number correlates with cell length but that chromosome duplication timing is asynchronous. Thus, while duplication is correlated to cell length, it is not coupled to cell division. Spatially, chromosome duplication occurs at random locations in the cell, but movement of each individual replisome remains confined after initiation. Together, these results elucidate chromosomal replication and segregation dynamics in a polyploid prokaryote.

Materials and Methods

Bacterial Strains and Growth Conditions

The wild-type *Synechococcus elongatus* PCC 7942 strain was acquired from the American Type Culture Collection (ATCC). *S. elongatus* cells were grown in solid BG11 medium following standard protocols with an illumination of 2000 lux at 30°C [14]. *S. elongatus* were transformed following standard protocols by incubating cells overnight in the dark with 100 ng of plasmid DNA and plating on selective media [15]. Antibiotics (kanamycin, spectinomycin, or chloramphenicol) were used at a concentration of 5 μg/ml. To prevent disruption of chromosome replication during growth, 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the media. Cells were then replica plated onto media with 50 μM IPTG [19] for visualization and further experiments.

Plasmid Construction

All cloning, unless otherwise stated, was done using a Biobrick-like strategy (SpeI as the upstream site and XbaI-HindIII-NotI as the downstream sites) [16]. 21 bp *lacO* operator sites were assembled with random ten bp spacers. *lacO* arrays were obtained from pLAU443 [17]. Two *lacO* arrays, with 120 *lacO* sites each, were then assembled with a kanamycin resistance marker inserted between them. Using Nhe1 and SalI restriction enzymes, this series of *lacO* arrays was then cloned into the neutral site 1 vector pAM2314 [18] or a vector containing homology regions to the terminus at 1.59 Mb in the *S. elongatus* chromosome. In the same vector LacI, fused to either the superfolder variant of green fluorescent protein (GFP) or yellow fluorescent protein (YFP) was inserted.

Image Acquisition and Analysis

Cells were plated onto BG11+50 μM IPTG [19] +2% agarose pads, which were transferred to a glass bottom dish (MatTek) for imaging. A Nikon TE-2000 microscope with a 100×1.4 numerical aperture objective equipped with an ORCA-ER CCD camera was used. Image acquisition utilized custom software, written using MATLAB (Mathworks), which interfaced with the microscope control package μManager [20]. Lighting necessary for cell growth during time lapse microscopy was controlled via a network AC power controller (IP Power 9258T), which also interfaced with MATLAB. Image analysis was performed using ImageJ [21], custom software written in MATLAB using the Image Processing Toolbox, and MicrobesTracker [22]. Cells were segmented using

phase contrast images and cell size was calculated. Chromosomes were identified as foci in fluorescent images and their location and number calculated. Tracking and segmentation were verified manually and corrected as necessary.

Single-Stranded-Binding (SSB) Protein Visualization

SSB protein genes were cloned from *S. elongatus*, fused to mOrange, and cloned into the neutral site 3 vector using methods described above. The resulting plasmid was transformed into cyanobacteria either alone or with the LacI-*lacO* plasmid and visualized using methods described above.

Fluorescently Labeled Nucleotides Incorporation and Imaging

Cells were grown in the presence of 0.3% pluronic F-68 to OD750 = 0.4. Pluronic F-68 concentration was then elevated to 3% and fluorescently labeled nucleotides tetramethylrhodamine-5–2′-deoxy-uridine-5′-triphosphate (Roche) were added at a final concentration of 3 μM. After growth to late log phase, cells were washed in PBS and imaged as described above.

Results

Fluorescent Tagging of the Genome Using a Repressor-operator System Reveals the Spatial Localization of Origins and Termini of Chromosomes *in vivo*

Organization of chromosomes has been studied in cyanobacteria using DAPI and fixed-cell staining methods [23,24]. However, these methods only give a static and low-resolution image of chromosome localization. In order to visualize and quantify chromosome dynamics *in vivo*, we used a fluorescent repressor-operator system (Fig. 1B). This system uses fluorescently-tagged DNA-binding proteins that bind to their cognate recognition sequences. Multiple proteins bound to operator arrays then appear as foci when imaged using fluorescence microscopy [25,26]. In our case, we used the LacI repressor fused to either yellow fluorescent protein (YFP) or the superfolder variant of the green fluorescent protein (GFP) as our DNA binding protein [27]. Simultaneously, an array of 240 *lacO* operator sites [1] was inserted using homologous recombination at various positions in the chromosome (Fig. 1B).

The precise location of the origin of replication (*oriC*) and the terminus region (*ter*) was predicted using the program Ori-Finder on the *S. elongatus* chromosome [28] and was recently confirmed with experimental data [29]. This analysis revealed that the origin of replication is located at the region defined as the start in the current chromosome sequence from NCBI. It is the intergenic region between *dnaN* and *ccbZp* and contains 11 *dnaA* boxes (consensus sequence *TTTTCCACA)* [28]. Interestingly, the *dnaA* gene, which is usually found near *oriC* in other species, was found elsewhere in the genome (1.1 Mb). The terminus region was not clearly defined via either GC skew or base disparity [28]. Due to the highly recombinant nature of the *S. elongatus* genome, the GC skew plot does not display a clear V-shape typical of organisms such as *E. coli* [30]. We reasoned that the terminus would be close to the region with the highest peak of GC disparity (Fig. S1, green line) so we inserted the *lacO* array at 1.59 Mb in the chromosome.

Visualization of tagged *oriC* showed multiple distinct foci throughout the cell (Fig. 1C). This is unlike previous observations of *oriC* localization in *E. coli* and *B. subtilis*, where *oriC* were replicated and maintained at the poles of cells [31,32]. Tagged termini also displayed multiple foci throughout the cell (Fig. 1D), whereas previously, termini in *E. coli* were found to migrate from poles to mid-cell during cell division [33]. Also, the origin and terminus in *E. coli* are spatially separate, confined to distinct regions of the cell [1]. We did not find such spatial specificity in *Syenchococcus* (Fig. 1D, bottom). Together, these data confirm earlier studies that *S. elongatus* does indeed have multiple copies of its chromosome throughout the cell cycle [34]. They also reveal that the localization of the *oriC* and terminus in *Syenchococcus* is distinct from that in other bacteria. The strains with the array integrated near the *oriC* showed a better signal-to-noise-ratio, so this strain was used in subsequent experiments.

Chromosome Duplication is Correlated to Cell Length and not Coupled to Cell Division

Using a custom made algorithm developed in MATLAB, we quantified the number of fluorescent foci (tagged *oriC*), representing the number of chromosome copies in each cell. The algorithm also allowed us to quantify cell boundaries and cell length using phase contrast images.

We found that the chromosome copy number distribution was not significantly different from a log-normal distribution ($n = 681$, χ^2 goodness of fit test, $h = 0$, $p = 0.2621$) with a mean of 4.62 copies per cell and a median and mode of 4 copies per cell (Fig. 2A). Cells harbored 1–10 chromosome copies. Interestingly, some cells contained an odd number of chromosomes and chromosome copy numbers other than 2^n copies. This observation does not fit the model typically observed in single replicon prokaryotes, where replication occurs synchronously [9]. Instead, it is similar to observations of asynchrony in *E. coli* replication mutants [35]. The chromosome copy numbers hint at asynchronous DNA replication, supporting previous fjndings of constant DNA synthesis rate over time; that is, DNA replication is not coupled to cell division [34].

We found that larger cells contained higher number of chromosome copies compared to smaller cells, strongly supporting a model where chromosomes replicate at a constant rate during growth. By observing chromosome numbers in growing cells, we found a linear correlation between cell length and number of chromosomes ($n = 660$, $r^2 = 0.57$, $p<0.001$) (Fig. 2B, red line). These results suggest that chromosome duplication is correlated to cell growth.

Replication Timing is Asynchronous: Only One Replisome is found in Most Cells at Any Given Time

In order to better understand the timing and spatial localization of single replication events *in vivo*, we fluorescently tagged single-stranded-binding (SSB) proteins. SSB proteins play a fundamental role during chromosome replication, coating single-stranded DNA that is temporarily exposed; thereby, preventing it from degradation [36]. Approximately 30 SSB proteins localize to the replisome in *E. coli*, and this method has been used extensively to track replisomes in other organisms [11,37].

Cyanobacterial SSB protein was fused to mOrange and expressed in cells with the origin of replication tagged with GFP (as described above). We found that SSB foci (SSB-mOrange) were co-localized with Origin-GFP foci present in the cell (Fig. 3A), indicating that tagging did not interfere with SSB function. SSB foci, therefore, correspond to active chromosome replication occurring in the cell. We found that at any given time, 85% of cells contain just one replisome, while 13.6% have two and only 1.3% have three (Fig. 3B). Since most cells contain only one actively replicating chromosome (SSB-mOrange foci) but more than one chromosome (Origin-GFP foci), our data show that, in most cells, only one copy of the chromosome is being replicated at any given

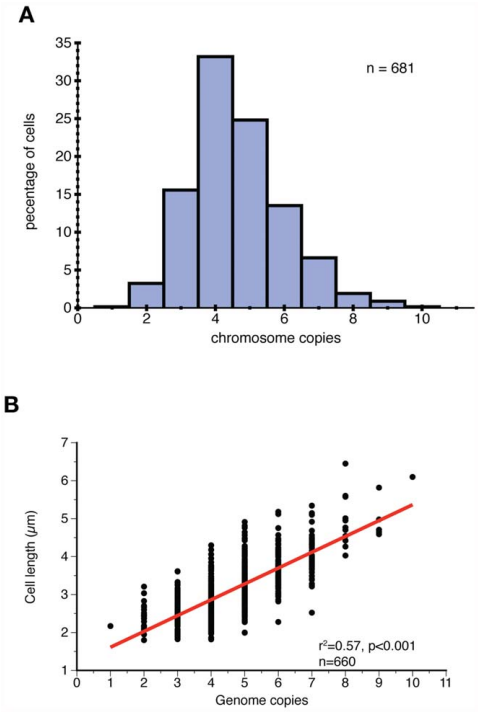

Figure 2. Chromosome duplication is correlated to cell length. (A) Distribution of chromosome number per cell is not significantly different from a log-normal distribution (n = 681, χ^2 goodness of fit test, h = 0, p = 0.2621). Most cells contain 4 chromosomes with values ranging from 1 to 10. (B) Chromosome copy number is correlated to cell length (n = 660, r = 0.7519, p<0.001), suggesting chromosome duplication is coupled to cell growth.
doi:10.1371/journal.pone.0047837.g002

time. This, along with chromosome number data (Fig. 2A) suggests that replication occurs asynchronously in *S. elongatus*.

Chromosome Duplication is Spatially Randomly Distributed

Since we found that cells containing multiple chromosomes typically replicate a single chromosome at a time (Fig. 3), we hypothesized that new chromosomes may be synthesized at a particular location in the cell, either at the poles or mid-cell. We investigated the spatial localization of the replication event and of newly synthesized chromosomes to determine if there is a spatial preference for replication.

To accomplish this, we segmented cells using a custom MATLAB algorithm and sub-segmented each cell along the major axis into 20 smaller regions. SSB foci location was distributed into these bins based on their distance from the pole of the cell along the major axis (Fig. 4A). This distance was normalized to the total length of the cell. We found that the distribution of SSB localization was not significantly different from a uniform distribution (Kolmogorov-Smirnov test, h = 0, p = 0.4867, k = 0.0456), suggesting random localization of replisomes. That is, duplication is equally likely to begin at any point along the length of the cell in the inner quintiles (20%–80%). SSB at the poles of cells were not included in this analysis, since nucleoid volume results in reduced probability of the replisome appearing at the edge of the cell. In addition, SSB foci are less likely to be found at the poles due to decreased cell volume at the ends. From this data, we found a striking absence of spatial preference for beginning replication.

In order to confirm our finding that chromosome replication occurs in random locations throughout the cell, we also fluorescently labeled newly synthesized DNA. To do so, we permeabilized cells using pluronic F-68 to allow uptake of fluorescently labeled nucleotides [38]. After cells grew to late log-phase, incorporation of fluorescent nucleotides into the chromosome was seen as foci. The images were automatically segmented and binned as described earlier. There seemed to be no spatial preference in newly synthesized DNA (Fig. 4B). Chromosome localization in the inner deciles (10%–90%) of the cell's major axis was not significantly different from a uniform distribution (KS test against uniform distribution. h = 0, p = 0.1027, k = 0.0579). We conclude that there is no spatial preference for chromosome duplication in the cell, confirming our previous results (Fig. 4a) from replisome localization data. These findings are in contrast to findings on chromosome replication in *E. coli*, which are preferentially localized to the center or the quartile points along the major axis (in cells with one and two replisomes respectively), as well as chromosome replication in other organisms, where chromosomes preferentially replicate at the poles or the middle of the cell [2,39,40].

Replication Fork Movement is Constrained

We concluded that replisomes appear at random locations throughout the cell, but how does their localization change over the course of a single duplication process? To investigate dynamic behavior of replisomes, time lapse imaging of SSB foci was used to track movement of replication forks. We followed replisomes over time, tracking individual foci every 5 seconds. We calculated a diffusion coefficient of $6.24\times10^{-5}\pm4\times10^{-5}$ $\mu m^2/s$ ($n = 20$) for replication forks, similar to the reported values for *E. coli* ($\sim10^{-4}$ $\mu m^2/s$) [11,41].

In order to determine whether movement is constrained over the time scale of a complete replication cycle, we tracked replisomes for one hour, acquiring images every two minutes (Fig. 4C). We found that replisomes were confined to a region of the cytoplasm, remaining close to their initial position. This behavior was observed in cells containing one or two replisomes, suggesting that this is not merely the result of observing two strands of the same chromosome. The small amount of movement over the time scale of replication suggests that replisome motion is constrained to the local cytoplasmic region in which it assembled. This is distinct from *E. coli* or *C. crescentus* in which replication follows directed motion between mid-cell and the edges of the cell. The data suggest that chromosomes are spooled through the replisomes rather than replisomes tracking along the chromosome. Together, these data give us an understanding of the spatial organization of chromosome duplication in *S. elongatus*: duplication starts in random locations throughout the cell and as duplication proceeds, the replisomes remain stationary.

Chromosomes Transiently Align During the Cell Cycle

The lack of spatial preference of duplication events supports the current model that organisms with multiple chromosome copies do not have an active chromosome segregation system [4]. It is currently believed that, just as in a high copy number plasmid

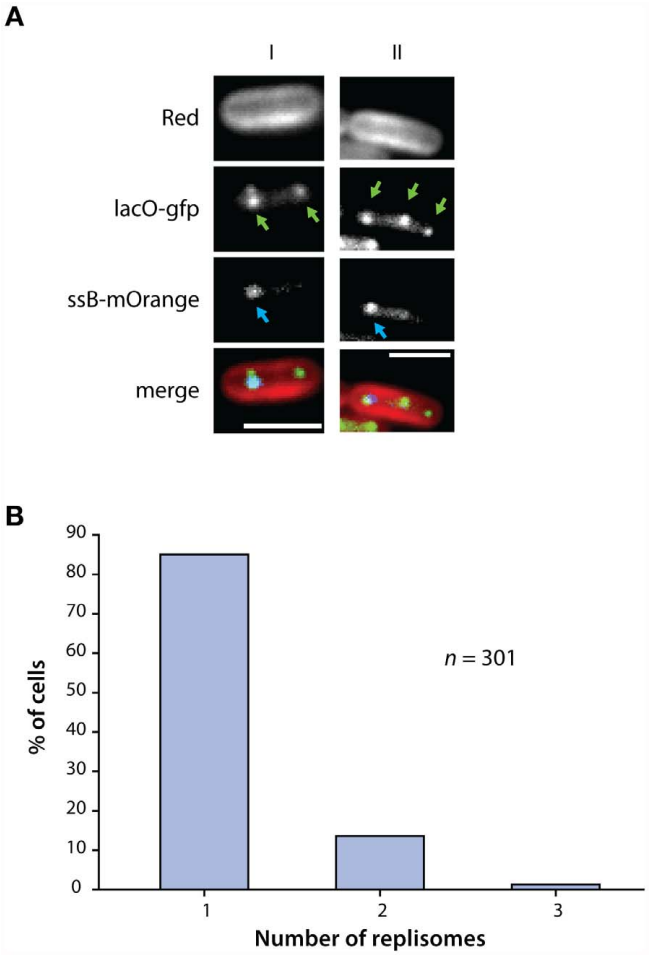

Figure 3. Chromosome duplication is asynchronous and is not coupled to cell division. (A) Single stranded binding (SSB) protein was tagged with mOrange. These were co-expressed in cells with LacI-GFP (NSI-*lacO*). Replisome localization appeared as foci and co-localized with tagged chromosomes (merge), indicating that tagging did not interfere with functioning of SSB. A cell with two chromosomes (left, green arrows) only contains one actively replicating chromosome (left, blue arrows). A cell with three chromosomes (right, green arrows) also only has one replisome (right, blue arrows). (B) Most cells contain one actively duplicating chromosome (85%), while the remaining contain two or three replisomes. Since most cells contain multiple chromosomes but only one replisome, this shows that chromosome duplication is asynchronous.
doi:10.1371/journal.pone.0047837.g003

system [13], a large number of chromosomal copies abolishes the need for an active segregation system since it is highly likely that each daughter cell obtains at least one copy by random. We surmised that our GFP-lacI-*lacO* method of visualizing chromosome dynamics *in vivo* would provide further insights into the dynamics of multiple chromosomes during segregation and allow us to follow up on previous works [24]. We found striking spatial organization, contrary to earlier models of random chromosome segregation in polyploid organisms.

Cells in a freely growing population displayed one of two phenotypes. Most of the cells displayed randomly localized chromosomes (85%, n = 289) (Fig. 5A right). However, some cells displayed chromosomes aligned along the major axis of the cell (Fig. 5A left). This surprising behavior prompted us to analyze the spatial arrangement of the chromosome copies over time. We tracked growing cells every hour for eight hours after which we could no longer distinguish signal from background (Fig. 5B). We found that chromosomes transiently aligned during the cell cycle: collapsing towards the middle and aligning evenly spaced along

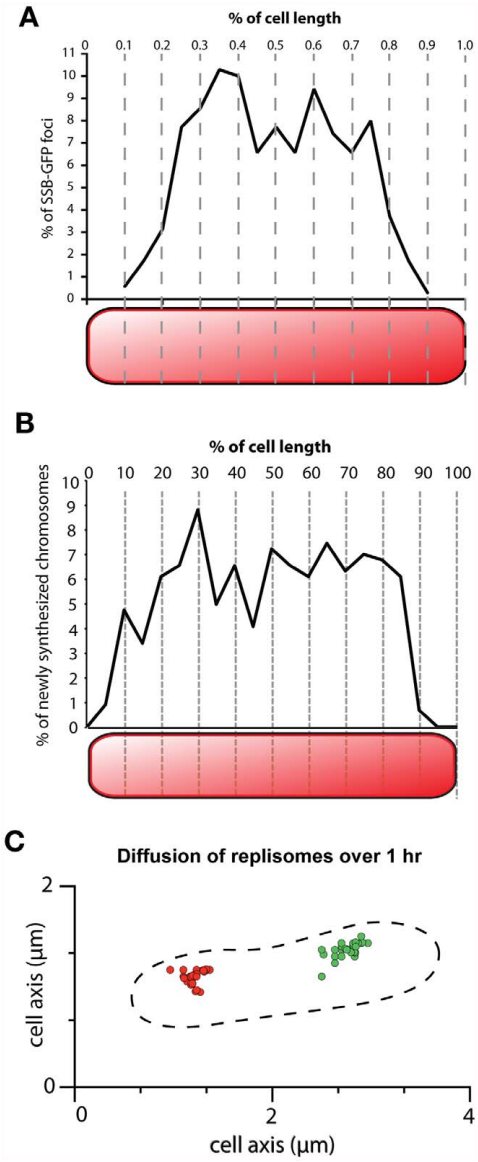

Figure 4. Replication occurs in a spatially random manner but replisomes undergo constrained movement. (A) Chromosome duplication events have no spatial preference within the cell. Cells (n = 297) were sub-segmented along the major axis into 20 smaller regions and SSB foci localization was binned. The distribution of SSB localization was not found to be significantly different from a uniform distribution (KS test, h = 0, p = 0.4867, k = 0.0456), suggesting spatially random replication events. Cell poles were excluded in this analysis because of boundary effects due to reduced cell volume and the volume of the nucleoid decreasing the likelihood of SSB foci found at the edges of rod shaped cells. (B) To confirm the random localization of chromosome duplication, newly synthesized chromosomes were visualized using fluorescent nucleotide incorporation. Fluorescent foci also show a uniform distribution (KS test, h = 0, p = 0.1027, k = 0.0579), confirming the results shown in (A). (C) Two replisomes (red and green) were tracked over the time scale of a complete replication cycle with images taken every two minutes. Replisomes show restricted occupation of domains in the cytoplasm, remaining close to their initial position, suggesting that chromosomes are spooled through the replisomes.
doi:10.1371/journal.pone.0047837.g004

the major axis of the cell (Fig. 5B, 4 hrs, yellow arrows). Shortly after (1 hr), spatial arrangement was lost. This transient spatial arrangement took place either one or two times during the acquisition window of 8 hours. In some of the cases, partial alignment was observed; that is, not all of the chromosomes aligned. This process was correlated to cell division, hinting that this process is mainly driven by the cell's commitment to division and may help maintain high fidelity chromosome segregation.

Chromosomes are Non-randomly Segregated

The transient alignment of chromosomes prompted us to hypothesize that chromosome segregation is not random, and that cells harboring the transient alignment have a high fidelity of chromosome segregation. This would be analogous to the spatial organization previously found regulating segregation of carboxysomes, microcompartments involved in the "carbon concentrating mechanism" in *S. elongatus* [42]. If true, chromosome segregation organization would be another example of order in bacteria, which were previously thought to be homogenous "bags of protein" with little to no internal organization [43]. To test our hypothesis, we assayed chromosome segregation by quantifying the number of chromosomes each daughter cell inherited after cytokinesis of mother cells (n = 48) that had 8 chromosome copies (Fig. 5C). We found a striking difference between the experimental results and a predicted binomial distribution based on random segregation to daughter cells (Fig. 5C Lilliefors test, $P<0.001$). Cells with 7, 9, and 10 chromosome copies were likewise found to undergo nonrandom segregation (Lilliefors test, $P<0.001$).

Discussion

The goal of this research was to understand the dynamics of chromosome replication and segregation in *S. elongatus*. To do so, we used the GFP-lacI-*lacO* fluorescent repressor-operator system to visualize chromosomes *in vivo*. Similar analyses have been performed in other bacterial species, including *E. coli*, and *B. subtilis* [1,44]. Here, however, we investigate live chromosomal dynamics in a bacterial species with multiple chromosomes. Having multiple copies of a single chromosome changes the parameters of the biological problem of replication and segregation that the organism must solve in order to successfully pass on its genetic information. The replication and segregation organization we found in *S. elongatus* differs from that in *E. coli* or other bacteria, and the difference can be attributed to the different challenges these bacteria face in passing on chromosomes to the next generation.

Bacterial chromosomal DNA is compacted into a nucleoid. In *E. coli*, the terminus and the origin can be overlapping or far apart from each other depending on the timing of replication and the cell cycle [45]. In our study, we labeled either the terminus or the origin of replication and analyzed origin movement in replication and segregation; however, different parts of the chromosome may be localized differently. We could not maintain the strain with the

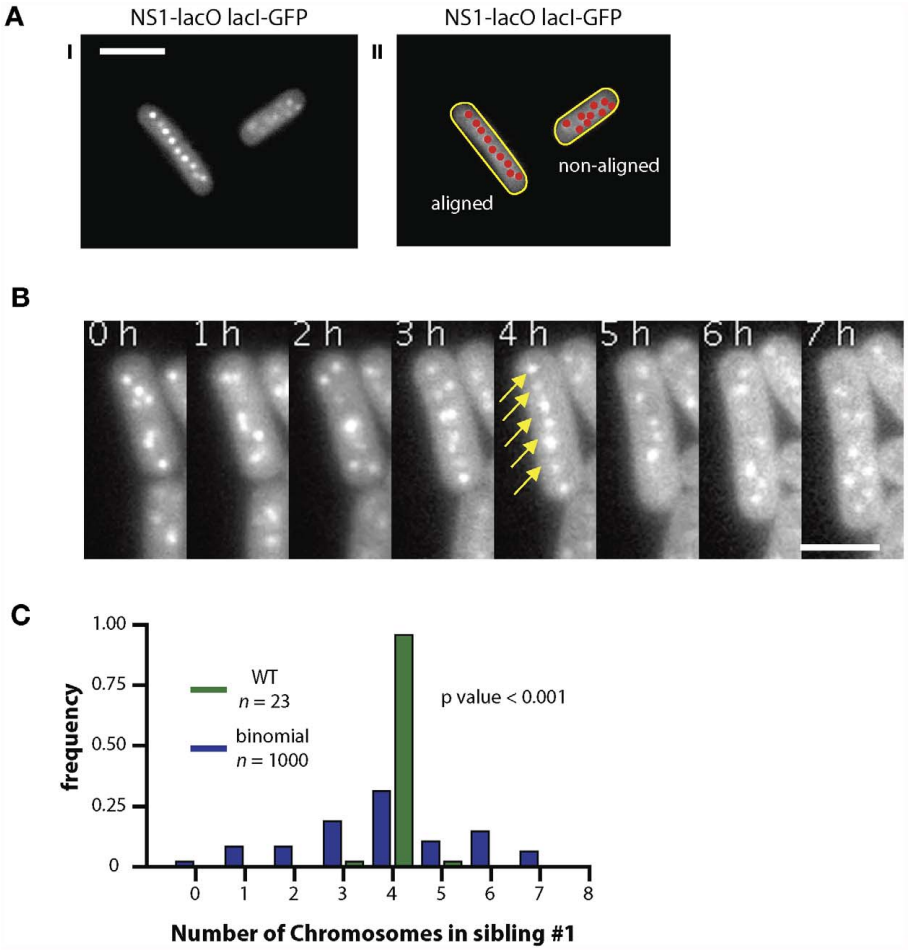

Figure 5. Chromosomes align transiently before non-random segregation occurs. (A) Chromosomes in cells visualized through LacI-GFP in an NSI-*lacO* background showed two different spatial arrangements. The chromosomes are either aligned along the major axis of the cell (I) or randomly localized (II). (B) Time-lapse imaging of single cells revealed a transient alignment of the chromosomes (4 hours, yellow arrows) approximately three hours before the cell entered cytokinesis. (C) Almost all sibling cells descended from mother cells containing eight chromosomes inherit four chromosomes, thus chromosome segregation is highly non-random.
doi:10.1371/journal.pone.0047837.g005

lacO array integrated at the terminus regions for long periods of time and we speculate that this may be due to a higher level of recombination occurring in that region of the chromosome. Integrations at other regions of the chromosome will clarify whether this is a global trend or specific to that region of the chromosome.

Our studies suggest that *S. elongatus* chromosome origins are randomly distributed throughout the cell. This is in contrast to specific localization of origins and termini in other bacterial species. In *E. coli* for example, the *oriC* is localized mid-cell and the terminus region is located at the poles of the cells [32]. In *C. crescentus*, both the *oriC* and *ter* are located at the poles [46]. In *V. cholerae*, one of its two origins is localized to the mid-cell, and the other is at the pole [47]. *B. subtilis* origins are located at the poles [31]. In these other species, replicated chromosomes initially occupy the same regions and require separation for proper segregation to occur. The relative positions of *S. elongatus* origin and terminus are still unclear. Two color experiments with tagging at both the terminus and origin (or other regions of the chromosomes) would help elucidate details of replication and segregation: how the DNA strand is situated within the cell during different phases of replication and segregation. It will also answer

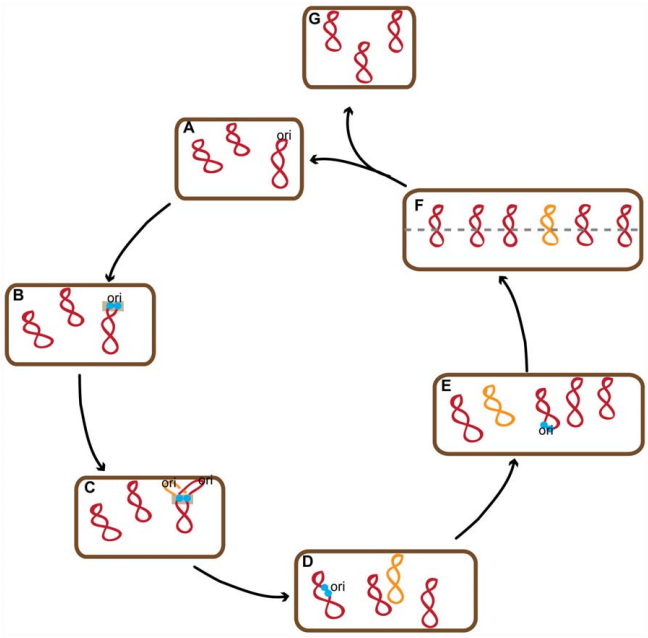

Figure 6. Model of chromosome replication and segregation in the polyploid bacterium *S. elongatus* PCC 7942. *S. elongatus* possess multiple copies of a single chromosome, shown in red (A). Chromosomes are duplicated asynchronously and coupled to cell growth (B, D, E). Newly synthesized chromosomes (orange) are synthesized in a spatially random manner (D,E,F). Replisomes (blue) assemble on a spatially random chromosome (B,D,E), but once initiated, their motion remains confined within the same region of the cell (grey box, B to C). Chromosomes transiently align (F) before non-random segregation and cytokinesis (F to G &A).
doi:10.1371/journal.pone.0047837.g006

how compact and region-excluded each chromosome is from its neighbors. The chromosomes may occupy distinct territories as has been previously observed in DAPI stained chromosomes [24]. The existence of local regions occupied by a single nucleoid can be advantageous for an organism with multiple chromosomes as it can undergo cytokinesis without the need to untangle chromosomes using FtsK or similar proteins.

Chromosome replication and cell division are fundamentally linked in most organisms [48], that is, replication must occur exactly once before (or during) cell division. Both events are believed to be regulated by mass doubling time, i.e. how long it takes for a cell to grow to double its mass [5]. Contrary to these studies in other organisms, asynchrony of chromosome replication has been observed in *S. elongatus* [8,29]. Indeed, recently, Watanabe et al. showed that cyanobacterial chromosomes replicate asynchronously based on mapping analysis and fixed cell staining methods [29]. They also suggested that while replication is still coupled to cell division (peaking at a few hours before cell division occurs), it is less stringently coupled than in *E. coli* or *B. subtilis*. In addition, cell division has been found to be gated by circadian rhythm [34], that is, cyanobacteria do not divide at night. Thus, timing of chromosome dynamics may be regulated by the circadian cycle. We found that replication in *S. elongatus* is independent of cell division and is instead correlated to cell length. We also found that chromosomes were replicated multiple times every cell division. Both replication and cell division may still be dependent on cell mass, but if so, the threshold of activation is much lower for chromosome replication than it is for cell division in *S. elongatus*. Observing chromosome copy number in cells where cytokinesis is inhibited may uncouple these dependencies. In addition, growing cells in multiple conditions leading to different cell division rates would also clarify the dependence between cell growth and chromosome copy number.

Several mechanisms ensure chromosomal replication takes place once every cell cycle and simultaneously from all origins in *E. coli* as well as in other prokaryotes [35,45]. These regulatory cellular processes give rise to a chromosome copy distribution where all possible values can be written in the form 2^n, where n is an integer. A Gaussian-like distribution of chromosome copy number similar to the one we observed has been found in *E. coli DnaA* mutants, which have disrupted initiation of replication [35,49]. Such a distribution could arise from the inability of some chromosomes to complete a single round of replication after the cell initiates synchronous replication. Alternatively, asynchronous initiation could also result in distributions containing copy numbers other than 2^n. We found that most cells only have one replisome and therefore only one actively replicating chromosome at a time. This supports asynchronous initiation as the mechanism that resulted in the observed Gaussian distribution of chromosome copy number.

There are two competing models for how replisomes proceed in DNA replication [37]. The first is the independent replisome

model, in which replisomes and replication forks track along the stationary chromosome. The second is the spooling replisome model, in which DNA is "spooled" through relatively stationary replisomes. A variation of the spooling model is the factory model in which the left and right replisomes are physically coupled throughout DNA replication. While early results favored a spooling model, recent results have shown that sister replisomes transiently separate. Another study observed that sister replisomes appear together early in S-phase but afterwards independently track DNA until they meet again as they near the terminus region. While our results suggest the spooling model is more likely in *S. elongatus*, we cannot rule out the independent model given the resolution limits of wide-field microscopy, especially if chromosomes each occupy exclusive local cytoplasmic regions during replication.

The spatial organization of chromosome origins along the major axis of the cell was observed for short periods (<1 hr) of the cell cycle. This spatial organization of chromosomes is surprisingly similar to eukaryotic mitosis, in which chromosomes align during metaphase before migration to the poles. However, in our current study, an active mechanism maintaining the organization was missing. Most other mechanisms of prokaryotic spatial organization studied such as carboxysome localization in *S. elongatus* [42] as well as low-copy plasmid segregation in *E. coli* [50,51], are maintained constantly throughout the cell cycle. We speculate that *S. elongatus* may not require an active segregation mechanism throughout the entire cell cycle in part because entropic forces may not be sufficiently disruptive to the arrangement of chromosomes post alignment and before cytokinesis. That is, due to having multiple chromosomes, *S. elongatus* do not have a stringent segregation problem and may not need constant organization. The transient alignment may enrich for rather than actively impose even segregation.

Conclusions

We have shown that the multiple copies of the chromosomes in *S. elongatus* can be tagged and tracked in living cells. Our model and findings are summarized in Fig. 6. Chromosomes are replicated in a linear-like fashion correlated with cell length in growing cells. By tracking replisomes *in vivo* we show that chromosome replication takes place in a confined region of the cytoplasm in accordance with the spooling replisome model and that chromosome replication does not happen preferentially in specific locations of the cell. Finally, we show *S. elongatus* segregates chromosomes to daughter cells in a non-random fashion, which we speculate may be the result of a cellular process that transiently organizes the chromosomes just before completion of cell division.

While this manuscript was in preparation, Jain et. al independently came to similar conclusions [52].

Supporting Information

Figure S1 GC disparity mapped the terminus region of the *S. elongatus* chromosome. Due to its highly recombinant nature, the genome of *S. elongatus* gives rise to a plot that does not display a clear V-shaped curve typical of organisms such as *E. coli*. We reasoned that the terminus region would be present within the vicinity of highest peak of GC disparity (green line) so we looked for a region amenable for integration and inserted a 240 repeat *lacO*-array in a region located at 1.59 Mb in the chromosome.
(TIF)

Author Contributions

Conceived and designed the experiments: BA AHC DFS PAS. Performed the experiments: BA AHC DFS. Analyzed the data: BA AHC DFS. Contributed reagents/materials/analysis tools: AHC BA DFS PAS. Wrote the paper: AHC BA.

References

1. Lau IF, Filipe SR, Søballe B, Økstad O-A, Barre F-X, et al. (2003) Spatial and temporal organization of replicating Escherichia coli chromosomes. Molecular microbiology 49: 731–743. Available: http://www.ncbi.nlm.nih.gov/pubmed/12864855. Accessed 2012 Jul 13.
2. Berlatzky IA, Rouvinski A, Ben-Yehuda S (2008) Spatial organization of a replicating bacterial chromosome. Proceedings of the National Academy of Sciences of the United States of America 105: 14136–14140. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2544591&tool=pmcentrez&rendertype=abstract. Accessed 2012 Jul 13.
3. Toro E, Shapiro L (2010) Bacterial chromosome organization and segregation. Cold Spring Harbor perspectives in biology 2: a000349. Available: http://cshperspectives.cshlp.org/content/2/2/a000349.full. Accessed 2012 Jul 16.
4. Graumann PL, Dame RT, Dorman CJ (2010) The Chromosome Segregation Machinery in Bacteria. Bacterial Chromatin (Google eBook). Springer. 31–48. Available: http://books.google.com/books?id=Esi4D4NjvskC&pgis=1. Accessed 2012 Jul 13.
5. Haeusser DP, Levin PA (2008) The great divide: coordinating cell cycle events during bacterial growth and division. Current opinion in microbiology 11: 94–99. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2397022&tool=pmcentrez&rendertype=abstract. Accessed 2012 Jul 13.
6. Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, et al. (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proceedings of the National Academy of Sciences of the United States of America 101: 9257–9262. Available: http://www.pnas.org/cgi/content/abstract/101/25/9257. Accessed 2012 Jul 16.
7. Griese M, Lange C, Soppa J (2011) Ploidy in cyanobacteria. FEMS Microbiology Letters 323: 124–131. Available: http://www.ncbi.nlm.nih.gov/pubmed/22092711. Accessed 2012 Nov 15.
8. Binder BJ, Chisholm SW (1990) Relationship between DNA cycle and growth rate in Synechococcus sp. strain PCC 6301. J Bacteriol 172: 2313–2319. Available: http://jb.asm.org/cgi/content/abstract/172/5/2313. Accessed 2012 Jul 15.
9. Mott ML, Berger JM (2007) DNA replication initiation: mechanisms and regulation in bacteria. Nature reviews Microbiology 5: 343–354. Available: http://dx.doi.org/10.1038/nrmicro1640. Accessed 2012 Jul 16.
10. Løbner-Olesen A, Skarstad K, Hansen FG, von Meyenburg K, Boye E (1989) The DnaA protein determines the initiation mass of Escherichia coli K-12. Cell 57: 881–889. Available: http://dx.doi.org/10.1016/0092-8674(89)90802-7. Accessed 2012 Jul 16.
11. Reyes-Lamothe R, Possoz C, Danilova O, Sherratt DJ (2008) Independent positioning and action of Escherichia coli replisomes in live cells. Cell 133: 90–102. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2288635&tool=pmcentrez&rendertype=abstract. Accessed 2012 Jul 13.
12. Jensen RB (2006) Coordination between chromosome replication, segregation, and cell division in Caulobacter crescentus. Journal of bacteriology 188: 2244–2253. Available:http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1428140&tool=pmcentrez&rendertype=abstract. Accessed 2012 Jul 13.
13. Ebersbach G, Gerdes K (2005) Plasmid segregation mechanisms. Annual review of genetics 39: 453–479. Available: http://www.annualreviews.org.ezp-prod1.hul.harvard.edu/doi/full/10.1146/annurev.genet.38.072902.091252. Accessed 2012 Jul 15.
14. Allen MM, Stanier RY (1968) Growth and Division of Some Unicellular Blue-green Algae. Journal of General Microbiology 51: 199–202. Available: http://mic.sgmjournals.org/cgi/content/abstract/51/2/199. Accessed 2012 Aug 12.
15. Clerico EM, Ditty JL, Golden SS (2007) Specialized techniques for site-directed mutagenesis in cyanobacteria. Methods in molecular biology (Clifton, NJ) 362: 155–171. Available: http://www.ncbi.nlm.nih.gov/pubmed/17417008. Accessed 2012 Jul 13.
16. Phillips I, Silver P (2006) A New Biobrick Assembly Strategy Designed for Facile Protein Engineering. Available: http://dspace.mit.edu/handle/1721.1/32535. Accessed 2012 Aug 12.
17. Lau IF, Filipe SR, Søballe B, Økstad O-A, Barre F-X, et al. (2004) Spatial and temporal organization of replicating Escherichia coli chromosomes. Molecular Microbiology 49: 731–743. Available: http://doi.wiley.com/10.1046/j.1365-2958.2003.03640.x. Accessed 2012 Jul 15.
18. Mackey SR, Ditty JL, Clerico EM, Golden SS (2007) Detection of rhythmic bioluminescence from luciferase reporters in cyanobacteria. Methods in

molecular biology (Clifton, NJ) 362: 115–129. Available: http://www.ncbi.nlm.nih.gov/pubmed/17417005. Accessed 2012 Aug 12.
19. Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, et al. (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proceedings of the National Academy of Sciences of the United States of America 101: 9257–9262. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 438963&tool = pmcentrez&rendertype = abstract. Accessed 2012 Jul 16.
20. Stuurman N, Amdodaj N, Vale R (2007) Micro-Manager: Open Source software for light microscope imaging. Microscopy Today 15: 42–43. Available: http://www.citeulike.org/group/8437/article/3968410. Accessed 2012 Aug 12.
21. Image processing with ImageJ (2004). Biophotonics international 11: 36–42. Available:http://igitur-archive.library.uu.nl/med/2011-0512-200507/UUindex.html. Accessed 12 August 2012.
22. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C (2011) High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. Molecular microbiology 80: 612–627. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 3090749&tool = pmcentrez&rendertype = abstract. Accessed 2012 Jul 15.
23. Hu B, Yang G, Zhao W, Zhang Y, Zhao J (2007) MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium Anabaena sp. PCC 7120. Molecular microbiology 63: 1640–1652. Available: http://www.ncbi.nlm.nih.gov/pubmed/17367385. Accessed 2012 Jul 13.
24. Schneider D, Fuhrmann E, Scholz I, Hess WR, Graumann PL (2007) Fluorescence staining of live cyanobacterial cells suggest non-stringent chromosome segregation and absence of a connection between cytoplasmic and thylakoid membranes. BMC cell biology 8: 39. Available: http://www.biomedcentral.com/1471-2121/8/39. Accessed 12012 Jul 13.
25. Gitai Z, Thanbichler M, Shapiro L (2005) The choreographed dynamics of bacterial chromosomes. Trends in microbiology 13: 221–228. Available: http://www.ncbi.nlm.nih.gov/pubmed/15866039. Accessed 2012 Jul 13.
26. Li Y, Sergueev K, Austin S (2002) The segregation of the Escherichia coli origin and terminus of replication. Molecular microbiology 46: 985–996. Available: http://www.ncbi.nlm.nih.gov/pubmed/12421305. Accessed 2012 Jul 13.
27. Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. Nature biotechnology 24: 79–88. Available: http://www.ncbi.nlm.nih.gov/pubmed/16369541. Accessed 2012 Jul 13.
28. Gao F, Zhang C-T (2008) Ori-Finder: a web-based system for finding oriCs in unannotated bacterial genomes. BMC bioinformatics 9: 79. Available: http://www.biomedcentral.com/1471-2105/9/79. Accessed 2012 Aug 6.
29. Watanabe S, Ohbayashi R, Shiwa Y, Noda A, Kanesaki Y, et al. (2012) Light-dependent and asynchronous replication of cyanobacterial multi-copy chromosomes. Molecular Microbiology 83: 856–865. Available: http://doi.wiley.com/10.1111/j.1365-2958.2012.07971.x. Accessed 2012 Jul 16.
30. Chang Y-F, Chang C-H (2011) CAGO: a software tool for dynamic visual comparison and correlation measurement of genome organization. PloS one 6: e27080. Available: http://dx.plos.org/10.1371/journal.pone.0027080. Accessed 2012 Aug 11.
31. Webb CD, Teleman A, Gordon S, Straight A, Belmont A, et al. (1997) Bipolar Localization of the Replication Origin Regions of Chromosomes in Vegetative and Sporulating Cells of B. subtilis. Cell 88: 667–674. Available: http://dx.doi.org/10.1016/S0092-8674(00)81909-1. Accessed 2012 Aug 6.
32. Gordon GS, Shivers RP, Wright A (2002) Polar localization of the Escherichia coli oriC region is independent of the site of replication initiation. Molecular microbiology 44: 501–507. Available: http://www.ncbi.nlm.nih.gov/pubmed/11972786. Accessed 2012 Aug 6.
33. Niki H, Hiraga S (1998) Polar localization of the replication origin and terminus in Escherichia coli nucleoids during chromosome partitioning. Genes & development 12: 1036–1045. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 316681&tool = pmcentrez&rendertype = abstract. Accessed 2012 Aug 6.
34. Mori T, Binder B, Johnson CH (1996) Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. Proceedings of the National Academy of Sciences of the United States of America 93: 10183–10188. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 38358&tool = pmcentrez&rendertype = abstract. Accessed 2012 Aug 6.
35. Skarstad K, von Meyenburg K, Hansen FG, Boye E (1988) Coordination of chromosome replication initiation in Escherichia coli: effects of different dnaA alleles. J Bacteriol 170: 852–858. Available: http://jb.asm.org/cgi/content/abstract/170/2/852. Accessed 2012 Jul 13.
36. Ha T, Kozlov AG, Lohman TM (2012) Single-Molecule Views of Protein Movement on Single-Stranded DNA. Available: http://www.annualreviews.org.ezp-prod1.hul.harvard.edu/doi/abs/10.1146/annurev-biophys-042910-155351. Accessed 2012 Aug 11.
37. Bates D (2008) The bacterial replisome: back on track? Molecular microbiology 69: 1341–1348. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2972702&tool = pmcentrez&rendertype = abstract. Accessed 2012 Jul 27.
38. Berlatzky IA, Rouvinski A, Ben-Yehuda S (2008) Spatial organization of a replicating bacterial chromosome. Proceedings of the National Academy of Sciences of the United States of America 105: 14136–14140. Available: http://www.pnas.org/cgi/content/abstract/105/37/14136. Accessed 2012 May 9.
39. Niki H, Yamaichi Y, Hiraga S (2000) Dynamic organization of chromosomal DNA in Escherichia coli. Genes & Dev 14: 212–223. Available: http://genesdev.cshlp.org/cgi/content/abstract/14/2/212. Accessed 2012 Aug 12.
40. Jensen RB, Wang SC, Shapiro L (2001) A moving DNA replication factory in Caulobacter crescentus. The EMBO journal 20: 4952–4963. Available: http://dx.doi.org/10.1093/emboj/20.17.4952. Accessed 2012 Jul 16.
41. Reyes-Lamothe R, Wang X, Sherratt D (2008) Escherichia coli and its chromosome. Trends in microbiology 16: 238–245. Available: http://www.ncbi.nlm.nih.gov/pubmed/18406139. Accessed 2012 Jul 13.
42. Savage DF, Afonso B, Chen AH, Silver PA (2010) Spatially ordered dynamics of the bacterial carbon fixation machinery. Science (New York, NY) 327: 1258–1261. Available: http://www.sciencemag.org/content/327/5970/1258.abstract. Accessed 2012 Mar 6.
43. Losick R, Shapiro L (1999) Changing Views on the Nature of the Bacterial Cell: from Biochemistry to Cytology. J Bacteriol 181: 4143–4145. Available: http://jb.asm.org. Accessed 2012 May 9.
44. Noirot-Gros M-F, Dervyn E, Wu LJ, Mervelet P, Errington J, et al. (2002) An expanded view of bacterial DNA replication. Proceedings of the National Academy of Sciences of the United States of America 99: 8342–8347. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 123069&tool = pmcentrez&rendertype = abstract. Accessed 2012 Jul 13.
45. Sherratt DJ (2003) Bacterial chromosome dynamics. Science (New York, NY) 301: 780–785. Available: http://www.sciencemag.org/content/301/5634/780.full. Accessed 2011 June 23.
46. Jensen RB, Shapiro L (1999) The Caulobacter crescentus smc gene is required for cell cycle progression and chromosome segregation. Proceedings of the National Academy of Sciences 96: 10661–10666. Available: http://www.pnas.org/cgi/content/abstract/96/19/10661. Accessed 2012 Aug 12.
47. Fogel MA, Waldor MK (2005) Distinct segregation dynamics of the two Vibrio cholerae chromosomes. Molecular microbiology 55: 125–136. Available: http://www.ncbi.nlm.nih.gov/pubmed/15612922. Accessed 2012 Aug 12.
48. Sherratt DJ (2003) Bacterial chromosome dynamics. Science (New York, NY) 301: 780–785. Available: http://www.sciencemag.org/content/301/5634/780.abstract. Accessed 2012 Jul 13.
49. Skarstad K, Boye E, Steen HB (1986) Timing of initiation of chromosome replication in individual Escherichia coli cells. The EMBO journal 5: 1711–1717. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1166998&tool = pmcentrez&rendertype = abstract. Accessed 2012 Jul 13.
50. Ringgaard S, van Zon J, Howard M, Gerdes K (2009) Movement and equipositioning of plasmids by ParA filament disassembly. Proceedings of the National Academy of Sciences of the United States of America 106: 19369–19374. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2775997&tool = pmcentrez&rendertype = abstract. Accessed 2012 Jul 13.
51. Gerdes K, Howard M, Szardenings F (2010) Pushing and pulling in prokaryotic DNA segregation. Cell 141: 927–942. Available: http://www.ncbi.nlm.nih.gov/pubmed/20550930. Accessed 2012 Jul 13.
52. Jain IH, Vijayan V, O'Shea EK (2012) Spatial ordering of chromosomes enhances the fidelity of chromosome partitioning in cyanobacteria. Proceedings of the National Academy of Sciences of the United States of America: 1211144109–. Available: http://www.pnas.org/cgi/content/abstract/1211144109v1. Accessed 2012 Aug 10.

E

Chen, A. H., Robinson-Mosher, A., Savage, D. F., Silver, P. A., Polka, J. K. (2013). The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. PloS One, 8(9), e76127.

OPEN ACCESS Freely available online

PLOS ONE

The Bacterial Carbon-Fixing Organelle Is Formed by Shell Envelopment of Preassembled Cargo

Anna H. Chen[1,2], Avi Robinson-Mosher[1,2], David F. Savage[3], Pamela A. Silver[1,2*], Jessica K. Polka[1,2]

1 Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts, United States of America, **3** Department of Molecular and Cell Biology, Department of Chemistry and Energy Biosciences Institute, University of California, Berkeley, California, United States of America

Abstract

Background: Cyanobacteria play a significant role in the global carbon cycle. In *Synechococcus elongatus*, the carbon-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) is concentrated into polyhedral, proteinaceous compartments called carboxysomes.

Methodology/Principal Findings: Using live cell fluorescence microscopy, we show that carboxysomes are first detected as small seeds of RuBisCO that colocalize with existing carboxysomes. These seeds contain little or no shell protein, but increase in RuBisCO content over several hours, during which time they are exposed to the solvent. The maturing seed is then enclosed by shell proteins, a rapid process that seals RuBisCO from the cytosol to establish a distinct, solvent-protected microenvironment that is oxidizing relative to the cytosol. These closure events can be spatially and temporally coincident with the appearance of a nascent daughter RuBisCO seed.

Conclusions/Significance: Carboxysomes assemble in a stepwise fashion, inside-to-outside, revealing that cargo is the principle organizer of this compartment's biogenesis. Our observations of the spatial relationship of seeds to previously formed carboxysomes lead us to propose a model for carboxysome replication via sequential fission, polymerization, and encapsulation of their internal cargo.

Citation: Chen AH, Robinson-Mosher A, Savage DF, Silver PA, Polka JK (2013) The Bacterial Carbon-Fixing Organelle Is Formed by Shell Envelopment of Preassembled Cargo. PLoS ONE 8(9): e76127. doi:10.1371/journal.pone.0076127

Editor: Michael T. Laub, HHMI, Massachusetts Institute of Technology, United States of America

Received August 21, 2013; **Accepted** August 23, 2013; **Published** September 4, 2013

Funding: This work was supported by an NIH (www.nih.gov/) Grant GM080177 and an NSF (www.nsf.gov/) Graduate Research Fellowship (to AHC), an NRSA postdoctoral fellowship (to ARM), and a Jane Coffin Childs (www.jccfund.org/) postdoctoral fellowship (to JKP), a Department of Defense Life Sciences Division of the US Army Research Office (http://www.arl.army.mil/www/default.cfm?page=200) Grant W911NF-09-0226 (to PAS) and funds from the Wyss Institute of Biologically Inspired Engineering at Harvard University (http://wyss.harvard.edu/). DFS was supported by a Department of Energy Office of Science Early Career Research Program (http://science.energy.gov/early-career/) (DE-SC0006394) through the Office of Basic Energy Sciences and the Energy Biosciences Institute at UC Berkeley. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: PAS serves on the Editorial Board of PLOS Biology.

* E-mail: pamela_silver@hms.harvard.edu

Introduction

Intracellular compartmentalization has long been considered the exclusive province of eukaryotes. However, prokaryotic cells also contain intracellular organelles, falling broadly into two categories. Some compartments are membrane-bound, including *Gemmata* nucleoids [1], cyanobacterial thylakoids [2], and magnetosomes [3]. Others are completely proteinaceous, such as gas vesicles [4] and metabolically active structures termed bacterial microcompartments. These form icosahedral structures that enclose enzymes required for certain metabolic processes, such as ethanolamine and propanediol utilization [5,6].

The carboxysome is one such microcompartment that encapsulates the carbon-fixing enzyme ribulose-1,5-bisphosphate carboxylase (RuBisCO) and carbonic anhydrase [7]. Carboxysomes are found in diverse cyanobacteria and chemoautotrophs and are crucial to the carbon sequestering capabilities of these organisms [8,9]. Inside the carboxysome, carbonic anhydrase converts bicarbonate to CO_2, which, along with ribulose-1,5-bisphosphate, is consumed by RuBisCO to produce 3-phosphoglycerate. Thus, the carboxysome serves to concentrate the metabolically inefficient RuBisCO enzyme and to increase the local concentration of CO_2. It has also been proposed that the carboxysome shell is selectively permeable to bicarbonate and ribulose-1,5-bisphophate while excluding oxygen, a competitor substrate of RuBisCO (Kinney et al. 2012). Finally, it has been speculated that the mature carboxysome must maintain a distinct internal oxidative

microenvironment to enable the enzymatic activity of carbonic anhydrase [10].

The mechanism and temporal sequence of carboxysome assembly is not known. The interior of the carboxysome is densely packed with its major cargo RuBisCO and a lower concentration of carbonic anhydrase. These are enclosed by proteins that form an icosahedral shell ~100nm in diameter [11,12]. Though their ultrastructural, but not phylogenetic, similarity to viral capsids may suggest that carboxyosomes assemble *de novo*, the mechanism of their biogenesis remains an unsolved problem [5].

There is evidence to suggest that shell proteins and cargo assemble together. Partially assembled carboxysomes have been observed by electron cryotomography, always containing both RuBisCO and shell proteins [13]. However, these data also argue that the cargo must have some intrinsic ability to self-assemble, as RuBisCO is seen to fill the inner layers of the nascent compartment. Indeed, *in vitro* evidence suggests that carboxysome contents can self-associate to form a structure without shell proteins [14]. Shell proteins of some carboxysomes can also independently assemble, forming empty microcompartments in the absence of cargo proteins [15].

The sequence by which these proteins assemble to form this complex organelle is not understood. We employ live cell fluorescence microscopy of *Synechcoccus elongatus* PCC 7942 to monitor the dynamics of carboxysome assembly. We find that carboxysomes originate near, and in some cases using material from, preexisting carboxysomes. They are born as small foci of RubisCO, which then grow over a period of hours. Shell proteins colocalize to these foci hours later, abruptly assembling to enclose the compartment and establish a protected internal microenvironment.

Results

Growing cells typically assemble one carboxysome at a time

Given existing structural evidence, we reasoned that solvent-accessible labeling strategies could be used to mark only RuBisCO inside carboxysomes in the process of forming. We constructed a strain of *S. elongatus* with SNAP labeled RuBisCO (RbcL-SNAP) expressed under an IPTG-inducible promoter (Figure 1A). After a 24 hour induction, we pulsed the cells with a fluorescent cell-permeable BG dye for 30 minutes and visualized them with fluorescence microscopy (Figure 1A-B). Among cells containing labeled foci, 88% had one, 10% had two, and 2% had three foci (Figure 1B, n=442). Thus, only a small subset of the average 3.7 carboxysomes per cell are labeled [16]. We used time-lapse microscopy to determine the difference between cells that contained labeled foci and those that did not. We found that all growing and dividing cells contained labeled foci (n=223), and those without labeled foci did not grow or divide. Thus, BG-labeled foci represent carboxysomes being assembling during the labeling pulse, mature carboxysomes are impermeable to the BG dye, and quiescent cells are not generating new carboxysomes. The distribution of foci numbers in growing cells indicates that *S. elongatus* carboxysomes are assembled one at a time rather than in parallel, as has been observed by electron microscopy in other genera of cyanobacteria [13].

New carboxysome are born colocalized with preexisting carboxysomes

Using time-lapse microscopy, we visualized the biogenesis of carboxysomes in live *S. elongatus* cells with green fluorescent protein labeled RuBisCO (RbcL-GFP) expressed under the IPTG-inducible promoter (Figure 1C). Fusions at this locus produce an additional 11% of wild-type levels of RbcL (Figure S1) and do not restrict growth [16]. We observed that new carboxysomes are formed at the site of preexisting carboxysomes. At the beginning of each birth event, a preexisting focus of RuBisCO is sometimes seen to take on an asymmetric character (Figure 1D, white arrow). Subsequently, a dimmer daughter focus emerges from the brighter mother carboxysome (Video S1).

The majority of new carboxysomes are generated at the site of preexisting carboxsyomes. Early after induction of RbcL-GFP, many unlabeled carboxysomes are still present in the cell, limiting our ability to determine whether all new carboxysomes colocalize with preexisting ones or arise at unrelated locations in the cell. To address this, we induced RbcL-GFP for 24 hours to label several carboxysomes in each cell, and then used particle tracking to generate lineage maps of carboxysomes (Figure 2). At the initiation of imaging, 65 “original” carboxysomes were present, and over the course of 26.1 hours, 106 new carboxysomes were formed. Of these, only two could not be assigned to visible mothers.

Furthermore, carboxysome birth events are spatially ordered, preferentially occurring at the quarter positions along the long axis of the cell (Figure 1J). These data are consistent with previous findings that cells have a mean of 3.7 carboxysomes positioned equally along their length by ParA (Savage, 2010). After birth, however, the new daughter carboxysome frequently localizes near the cell pole (Figure 1A and 1D). Quantification of RbcL-GFP intensity reveals that birth events are highly asymmetric, with an average daughter–mother intensity ratio of ~1:4 (n=141, Figure 1K). We also observe that some birth events are correlated with rapid motions of either mother or daughter or both (Video S1). Indeed, automated tracking of carboxysome velocities suggests that carboxysome velocity is variable, with the mean maximum speed of a carboxysome being over 60nm per minute (Figure S2C). This may be related to the approximately 100nm per minute movement of ParA, assuming a 3μm cell [16]. Analysis of individual tracks reveals that mean carboxysome velocity is higher in the first several hours after birth (Figure S2F).

Elongated bar carboxysomes divide in two. These carboxysomes are well-documented by electron microscopy and are found in normal *Synechococcus* cells [13,17], though higher frequencies (up to 20% of all carboxysomes) are associated with environmental carbon limitation or mutations that compromise carbon fixation [14,18]. Their elongated morphology (1-3μm in length) provides a means to study the spatial organization of carboxysomes above the resolution limit of light microscopy. In our RbcL-GFP strain 0.5% of labeled

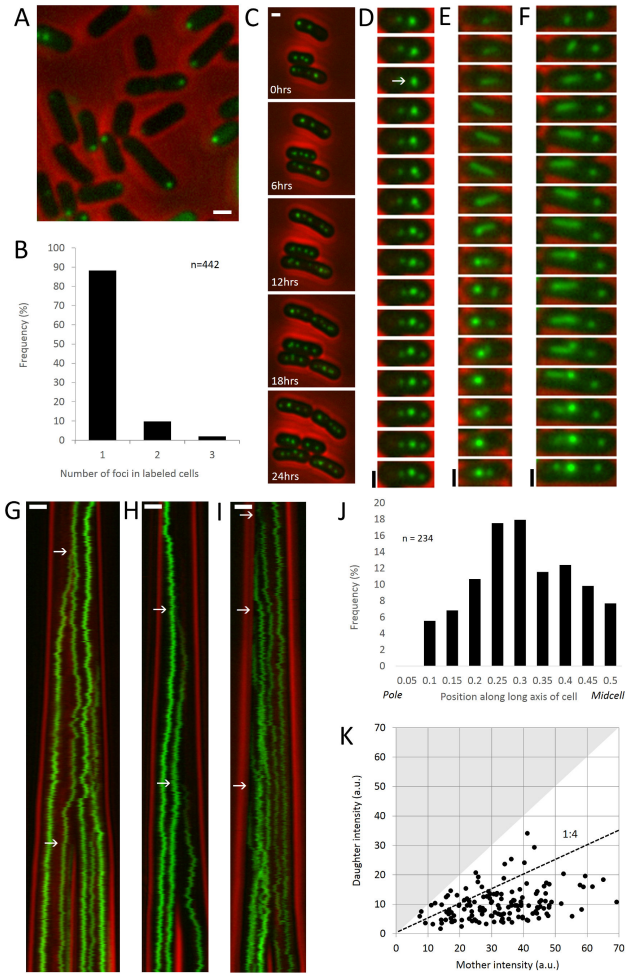

Figure 1. Carboxysomes are born one at a time at the site of preexisting carboxysomes. (A) In pulse-chase labeling of RbcL-SNAP in live *S. elongatus* cells, actively assembling carboxysomes with solvent accessible RbcL-SNAP are labeled with BG dye. Red: phase contrast. Green: RbcL. Scale bar: 1μm. (B) The distribution of the number of SNAP labeled carboxysomes, indicating active assembly, in cells directly after labeling (n=442). (C) The biogenesis of carboxysomes can be monitored from long timelapses. Red: phase contrast. Green: RbcL-GFP. Scale bar: 1μm. (D) Montage showing the formation of new carboxysome at the site of a preexisting carboxysome. White arrow indicates the birth event. Panel height: 25 pixels. Time interval: 3 minutes. (E–F) RuBisCO foci elongate into bar carboxysomes that subsequently split into two carboxysomes. Scale bar: 1μm. Time interval: 75 minutes. (G–I) Kymographs of RbcL-GFP in growing and dividing cells. Carboxysome birth events are indicated by white arrows. Scale bar: 1μm. Time interval: 3 minutes. (J) Spatial distribution of 234 birth events along the long axis of the cell. Quarter cell positions are favored. (K) Relative intensity of 141 pairs of new (daughter) carboxysomes and the preexisting carboxysomes to which they initially colocalize (mothers) reveals that birth events are highly asymmetric, with mean daughter intensity being 1/4 that of the mother. Because pairs are sorted into dim (daughter) and bright (mother) pairs, no data points can fall into the shaded area. Dotted line indicates a 1:4 ratio.

doi: 10.1371/journal.pone.0076127.g001

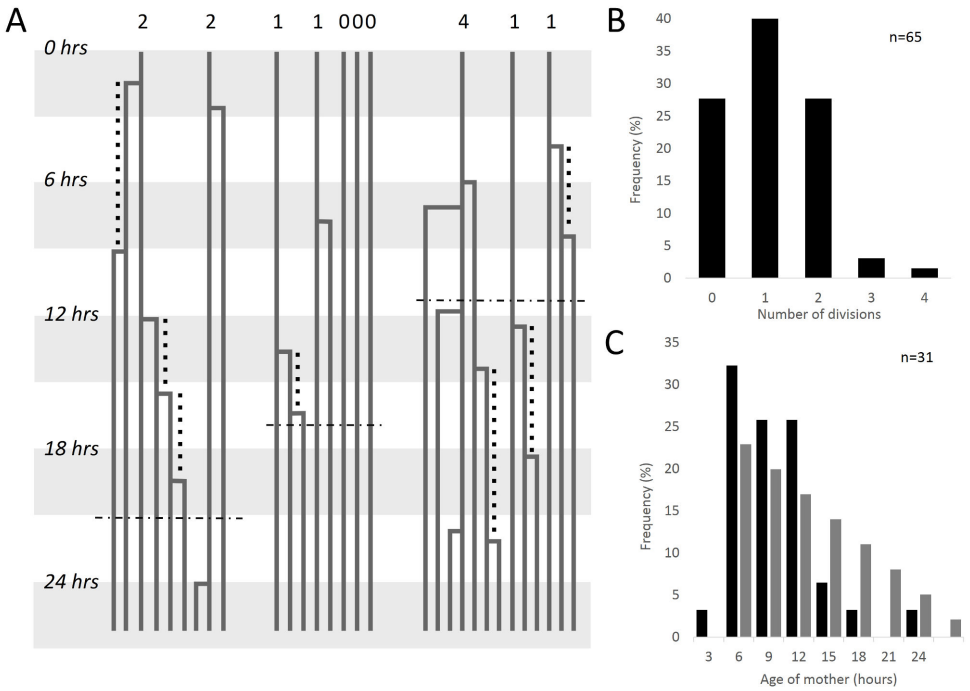

Figure 2. Mapping of carboxysome lineages reveals that new organelles undergo an initial refractory period before producing daughters of their own. (A) Example lineages of carboxysomes from three out of 25 total cells analyzed from a 522 frame movie taken at 3 minute intervals over approximately 26.1 hours. Each line represents a carboxysome tracked through time, with right-angle connectors joining daughters to mothers. Digits at the top of the panel indicate the number of times carboxysomes present at the beginning of the movie have colocalized birth events over the course of the analysis (26.1 hours), represented in the histogram in panel B. Vertical dotted lines indicate the measurable age of mothers when a daughter appears, represented in the histogram in panel C. Horizontal dotted lines indicate the time of cell division. (B) Histogram of the number of births colocalized to original carboxysome in the entire dataset (n = 65). (C) Histogram of measurable ages of mothers tabulated over the entire dataset (n = 31).

doi: 10.1371/journal.pone.0076127.g002

carboxysomes are bars after induction with 25μM IPTG (n=191 total), and 1% are bars after 50μM IPTG (n=395 total). These carboxysomes colocalize with shell protein, though their redox state suggests that they are immature (Figure S3, compare to Figures 3 and 4). Bar carboxysomes begin as puncta that elongate and subsequently collapse or split into two carboxysomes (Figure 1E-F and Video S1).

Maternal age influences the frequency of carboxysome births

By further analyzing the lineage maps, we found that new carboxysomes are more likely to be born near recently formed carboxysomes than near older ones. For each birth event, we measured the age of the mother carboxysome if it was born during the course of our observations (dotted lines in Figure 2A and histogram in Figure 2C). Strikingly, after a carboxysome is born, there is a marked refractory period until a new colocalizing daughter appears. This is characterized by a lack of births in the first three hours of its lifetime and suggests the structure must mature before another birth event occurs. Immediately following this refractory period, there is a burst of birth events. However, our imaging interval favors the observation of early birth events over late ones. To determine whether this apparent burst is significant, we compared our data to a model where birth probability is constant regardless of carboxysome age, adjusting for the limitations of our imaging interval. Our observed distribution of birth ages (black bars, Figure 2C) is significantly different from the theoretical

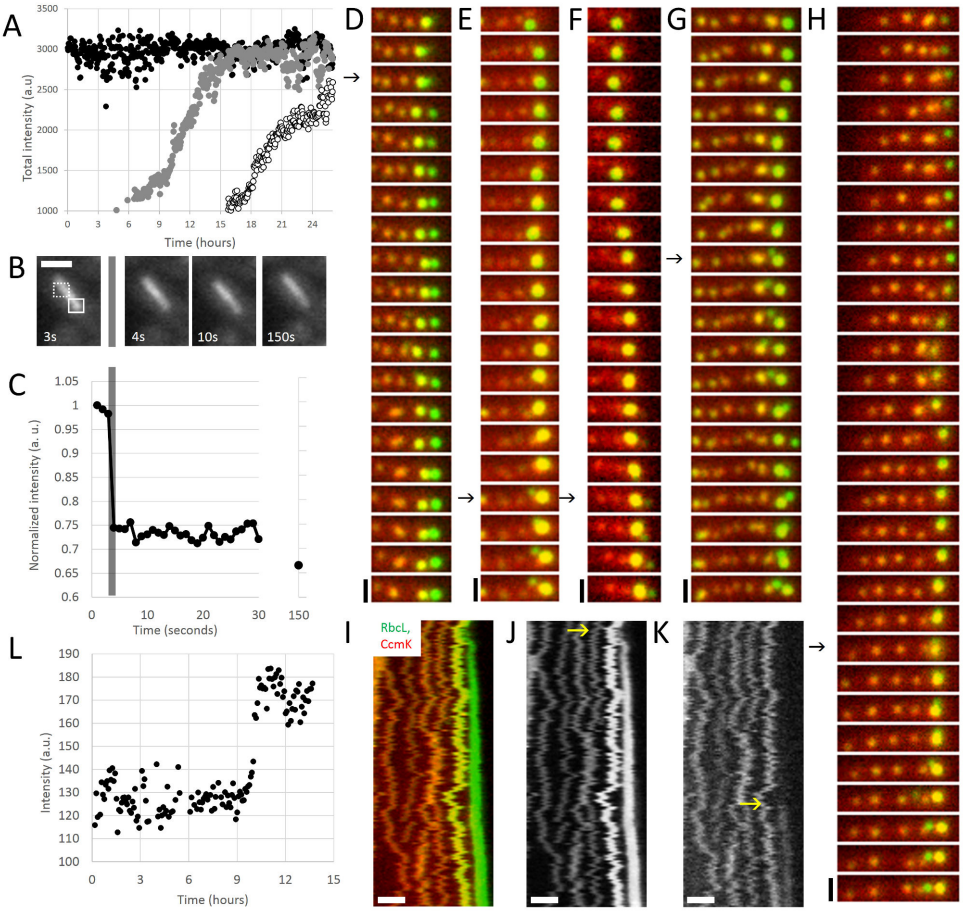

Figure 3. RuBisCO slowly forms a structured assembly prior to rapid colocalization of shell protein. (A) RuBisCO assembly, as measured by fluorescence intensity, follows sigmoidal kinetics. Each trace represents a new carboxysome. Cell is same as that depicted in Figure 1I. Imaging interval: 3 minutes. (B) Fluorescence recovery after photobleaching of a segment of a bar carboxysome. Solid box shows bleached area. Unbleached area (dashed box) was used for photobleaching correction. Cells were imaged at regular intervals after bleaching to assay for recovery. Scale bar: 1μm. (C) Quantification of FRAP in (B). Grey bar indicates bleaching event, when fluorescence sharply decreases. No recovery was seen after 150 seconds. (D–H) Time lapse of RbcL-mOrange (green) and CcmK4-GFP (red). Arrows indicate birth events of carboxysomes. Newly born RuBisCO initially buds off without shell protein. Shell protein colocalizes to RbcL-GFP foci hours after birth. In some cases (G and H), shell protein assembly is correlated with the formation of a new RuBisCO focus. Scale bar: 1μm. Time interval: 25 minutes. (I–K) Kymograph of RbcL-mOrange (J) and ccmK4-GFP (K) assembly. Shell protein assembly (yellow arrow in K) initiates well after RuBisCO birth event (yellow arrow in J). Scale bar: 1μm. Time interval: 5 minutes. (L) Individual trace of the fluorescence intensity of a CcmK4 focus in the process of formation. Time interval: 5 minutes.

doi: 10.1371/journal.pone.0076127.g003

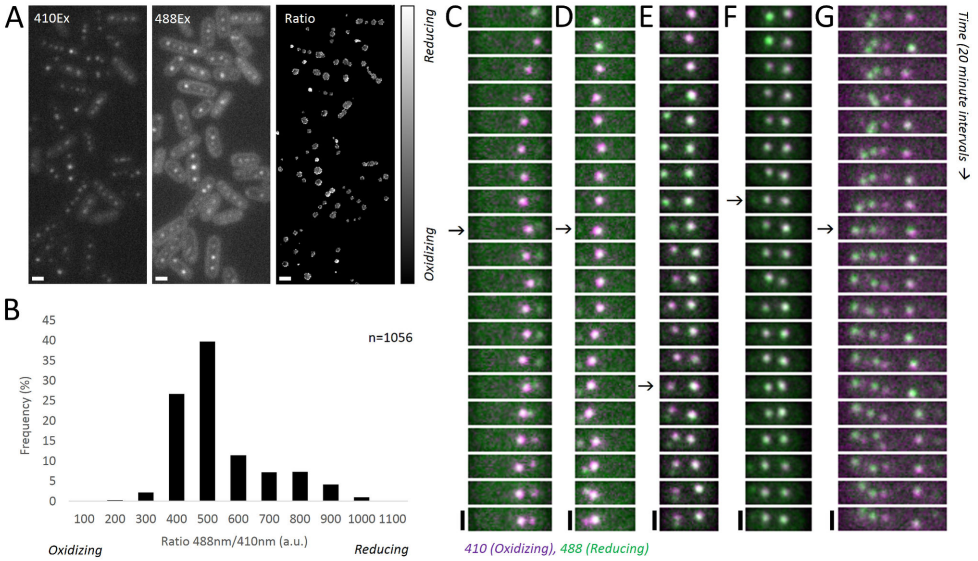

Figure 4. The carboxysome oxidizes over the course of its maturation. (A) RbcL-roGFP1 excited with 410nm (left) and 488nm (middle) produces ratiometric (488nm/410nm) differences in emission (right). Scale bar: 1μm. (B) A histogram of this ratio measured at each carboxysome focus reveals an asymmetric distribution biased toward a relatively oxidized state. (C–F) Montages of RbcL-roGFP1 show transitions from predominantly 488nm excitation (green) to 410nm excitation (magenta) over the maturation period of carboxysomes. Carboxysomes establish an oxidizing state before the appearance of a new carboxysome, rarely reopening to the cytosol after an initial closure (G). Arrows indicate birth events. Scale bar: 1μm. Interval: 20 minutes.

doi: 10.1371/journal.pone.0076127.g004

distribution predicted by the age-independent model (grey bars, Figure 2C) (Kolmogorov-Smirnov test, h=1, p-value = 0.0416, k=0.2438). In comparison, our data show that birth rates are enriched in the first 12 hours of the lifetime of the mother. We measured the birth rate in the 9 hours following this three hour refractory period at 0.42 per carboxysome (n=57 carboxysomes), versus 0.28 per carboxysome over 9 hours for those at least 12 hours old (n=65).

While young carboxysomes have colocalized birth events more frequently during the burst period, mature carboxysomes have daughters randomly. By tabulating the number of times that carboxysomes visible at the start of the time lapse colocalized with new birth events, we found that this distribution of events is nearly Poissonian, with a mean of 1.1 and a variance of 0.8 (Figure 2B). This indicates that births near mature mothers is a random process. Furthermore, the probability that a preexisting carboxysome has daughters in the second half of the movie is not influenced by whether it did (0.36) or did not (0.37) in the first half of the movie, suggesting that births near mature carboxysomes are independent events. Furthermore, timing of carboxysome birth events is not correlated to cell divisions (horizontal dotted lines in Figure 2A). The distribution of birth event timing as a fraction of the cell cycle is not significantly different from a random uniform distribution across the cell cycle (Kolmogorov-Smirnov test, h=0, p=0.3222, kstat=0.13).

Carboxysome biogenesis begins with a sigmoidal assembly of RuBisCO

To understand the nature of the maturation process, we followed the assembly of RuBisCO and coat protein (CcmK4). We first examined the kinetics of RuBisCO assembly by measuring the intensity of RbcL-GFP foci over time. This indicated that RuBisCO assembles over the course of many hours in distinct phases that display sigmoidal kinetics (Figure 3A). While individual carboxysomes assemble at different rates, we observed three regimes: a lag phase, followed by rapid assembly, and finally a plateau phase - assembly kinetics reminiscent of nucleation condensation polymers.

Once assembled, RuBisCO does not freely diffuse inside carboxysome foci. To probe the nature of assembled RuBisCO-GFP, we monitored fluorescence recovery after photobleaching (FRAP) of bar carboxysomes (Figure 3B-C and Figure S4). Bar carboxysomes are sufficiently large such that only a segment of the bar was bleached (Figure 3B solid box) while the rest of the bar remained fluorescent (Figure 3B

dashed box). No recovery of fluorescence was seen up to 150 seconds after bleaching (Figure 3C). RuBisCO hexadecamers are roughly 500kDa in size, and freely diffusing protein complexes of similar molecular weight have been reported to recover in *in vivo* FRAP experiments in less than two seconds [19]. This discrepancy indicates that assembled RuBisCO does not freely exchange with monomers in the cytoplasm or in the rest of the carboxysome; rather, assembled cargo is static on the timescale of minutes.

Shell proteins rapidly colocalize with RuBisCO late in the assembly process

To determine the relative kinetics of RuBisCO and shell assembly, we performed time-lapse microscopy of cells expressing inducible RuBisCO fused to mOrange and the shell protein CcmK4 fused to GFP under a constitutive promoter. At the beginning of the observation interval, recently born carboxysomes show strong signal in the RuBisCO channel, while old carboxysomes show weak, background levels of signal. New carboxysomes begin assembly with little to no detectable shell protein (Figure 3D–H and Video S2). Instead, shell protein associates with nascent carboxysomes at a later point: the mean time between the first appearance of RuBisCO and detectable shell protein was 4.7 hours (+/-2.2 hours, n=54). Kymographs of the formation process are shown in Figure 3I–K, where the shell protein suddenly colocalizes 8 hours after the birth event of the RbcL-GFP focus. The assembly of shell protein completes rapidly in contrast to the many hours required for RuBisCO assembly (Figure 3A); shell intensity reaches steady state in less than two hours (Figure 3L and Figure S5). Interestingly, in a fraction of cases, we observed the apparent birth of small daughter focus from a carboxysome 3.1 hours (+/- 1.1 hours, n=9) after detectable shell colocalized with the mother (Figure 3D–H). The timing of these events correlates with the burst of births following a maternal refractory period in lineage maps (Figure 2C).

The carboxysome establishes a unique microenvironment late in the assembly process

The enzymatic activity of carbonic anhydrase relies on an environment more oxidative than the bacterial cytosol [20]. This predicts that the carboxysome must maintain an internal oxidizing state. To monitor changes in the redox state of the carboxysome over time, we tagged RuBisCO with the redox-sensitive roGFP1 [21]. The excitation spectrum of this protein shifts from one dominated by a maxima at ~488nm under reducing conditions to one dominated by a maxima at ~410nm under oxidizing conditions. By measuring the ratio of these two channels, we observed that carboxysomes display varying redox states within the same cell (Figure 4A). Over the entire population, we find that carboxysomes are distributed across a range of redox states (Figure 4B) but that the distribution is skewed toward oxidizing states (median=440, mean=487).

The late assembly of shell proteins on nascent carboxysomes predicts that maturing foci of RuBisCO share the reducing cytosolic environment. Indeed, monitoring changes in redox state revealed that newly formed RuBisCO foci are relatively reduced, regardless of whether imaging is started 6 (Figure 4C-D) or 24 (Figure 4E–G) hours after induction. As the carboxysome matures, RuBisCO-roGFP1 oxidizes, indicating the establishment of a distinct microenvironment (Video S3). Though the low signal from roGFP1 prohibits the time resolution required to measure the carboxysome lineage, the most recently synthesized carboxysome typically oxidizes before a new one appears (Figure 4C–F). Rarely, a carboxysome is apparently born from a mother rapidly switching between oxidized and reduced states (Figure 4G). Taken together, our data show that new carboxysomes are born concomitant with shell closure and establishment of the oxidizing microenvironment in the previously synthesized carboxysome.

Carboxysomes persist over the cell cycle and their cargoes can be redistributed to daughter carboxysomes

We used the BG pulse-chase experiments (Figure 1A-B) to track the lifetime and fate of RuBisCO assemblies over days. We observed that labeled RuBisCO from one initial focus can partition into two or more daughter carboxysomes, and that it persists over the time interval of the experiment (45hrs) (Figure 5A). In some cases, the intensity of the mother carboxysome could be seen to decrease with the birth of a new focus (Figure 5B), indicating repartitioning of RuBisCO to new carboxysomes from old ones. In other cases, no decrease in mother intensity was detectable with the appearance of other BG foci. This is perhaps due to either splitting events with signal changes beneath our detection limit or the assembly of residual labeled cytosolic RuBisCO (Figure 5B, grey trace). During the period of this pulse-chase experiment, the number of labeled foci per cell was either invariant or increased (Figure 5C). Interestingly, disappearance of foci was never observed, suggesting that carboxysomes are not degraded, but rather passed down through generations.

Discussion

We show that the *in vivo* biogenesis of carboxysomes occurs by preferential assembly on preexisting RuBisCO structures that later separate from mother carboxysomes. These stable, cytosol-accessible nuclei grow over a period of hours until shell proteins abruptly enclose the carboxysome, establishing a microenvironment distinct from the cytosol. This maturation can be coincident with the release of a new colocalizing daughter seed of RuBisCO.

Carboxysomes are the major carbon-fixing centers of the photosynthetic cyanobacterium *S. elongatus*; thus, maintaining an appropriate number of organelles is vital to the cell [16]. The assembly of these compartments is regulated in two ways: 1) by formation of one carboxysome at a time, and 2) by regulation of their geometry. Our data suggest that both of these constraints arise from the assembly properties of the components.

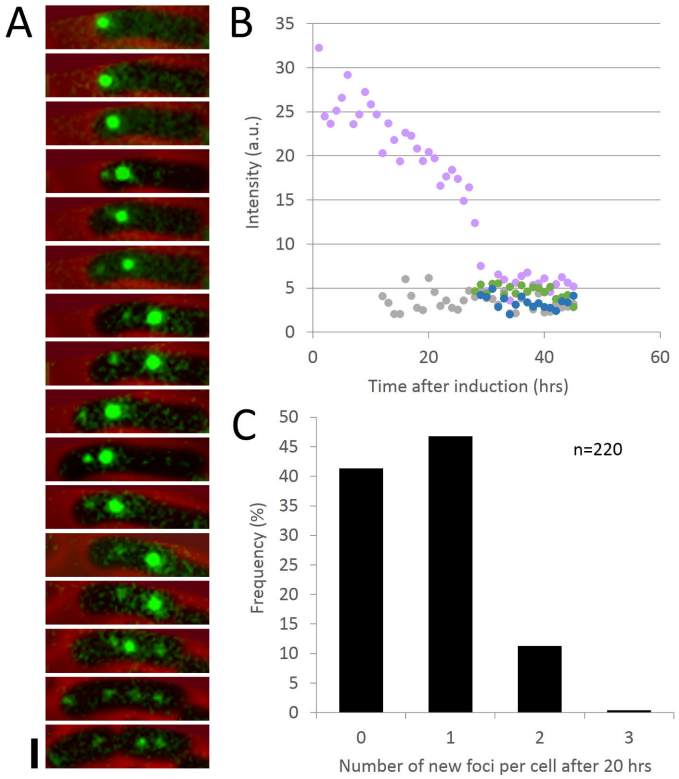

Figure 5. A solvent-accessible dye pulse labels foci that subsequently divide, but do not dissipate. (A) Montage showing one labeled RuBisCO focus partitioning into two or more daughter carboxysomes and persisting over the time interval of the experiment. Red: phase contrast. Green: RbcL-SNAP. Scale bar: 1μm. Time interval: 2 hours. (B) Intensity of the mother carboxysome (pink trace) sometimes decreases when new daughters (green and blue traces) are born. In other cases, the decrease is not detectable (grey trace). (C) Distribution of the number of new carboxysome foci formed per cell over the course of an experiment (20 hours). RuBisCO foci either persisted or divided over 20 hours (n = 220). All original foci were detectable at the end of the experiment.

doi: 10.1371/journal.pone.0076127.g005

Cargo assembly is the primary organizer of carboxysome biogenesis

Carboxysome biogenesis is tuned to produce one structure at a time (Figure 1B). This may be an energetically efficient strategy, as focused assembly minimizes the net time carboxysomes spend in an incomplete state, during which they cannot deliver energetic benefits to the cell. In order to achieve focused assembly, there must be a kinetic barrier to spontaneous nucleation of cargo so that growth occurs only on preformed seeds.

Several lines of evidence support a nucleation-limited assembly mechanism of RuBisCO. First, the one-at-a-time assembly process suggests that templated assembly is favored over *de novo* nucleation. Second, the sigmoidal kinetics of RuBisCO assembly are reminiscent of nucleation-limited polymers. Third, the elongation of some RuBisCO seeds into bar carboxysomes suggests that this assembly is an extensible process, capable of producing structures far larger than mature icosahedral carboxysomes. Fourth, FRAP of bar carboxysomes demonstrates that assembled RuBisCO does not freely diffuse, again reminiscent of a polymer lattice with stabilizing interactions between neighboring subunits.

Previous work also supports this idea. Contents from purified carboxysomes can self-assemble in a concentration-dependent

manner *in vitro* [14]. Furthermore, RuBisCO inside carboxysomes is organized into a lattice [13,22,23] implying that multiple self-associating interactions direct cargo to fill the interior layers of the carboxysome.

Shell assembly specifies organelle size and limits further addition

Polymerized cargo appears to be stable and capable of extending far beyond the geometry of a mature icosahedral carboxysome, as suggested by the existence of bar carboxysomes. However, most mature carboxysomes are homogeneous in size. We propose that the rapid enclosure by the shell protein not only limits further cargo assembly by isolating the assembled RuBisCO from the cytosolic pool of subunits, but also sets the size of the carboxysome. Our data suggest two mechanisms for the size determination of RuBisCO assemblies: 1) size-selective enclosure, and 2) bisection of excess cargo.

We observe the assembly of shell protein only late in the biogenesis process (Figure 3I–K), presumably when the RuBisCO lattice reaches a given size. The topology of the growing RuBisCO seed thus would present a multivalent binding surface, with curvature depending on the size of the overall assembly. It is known that shell protein also self-associates into structures of a given radius [15]. Therefore, we speculate that when the curvature of the RuBisCO assembly matches that of the shell, RuBisCO-shell interactions organize shell-shell interactions, facilitating the assembly process. In other words, the intrinsic structure of the shell may act as a topological sensor that regulates timing of RuBisCO enclosure, ensuring that nascent carboxysomes reach a minimum size before encapsulation.

We also speculate that the polymerization of the shell can bisect a RuBisCO assembly to generate a mature carboxysome and a new RuBisCo seed (Figure 6C and D). This shell-mediated pinching hypothesis presents a parsimonious explanation for the increased birth rates from young mothers (Figure 2) and the coincidence of new seed formation with both shell association and the establishment of a distinct microenvironment in the mother (Figures 3 and 4). Our data also support non-pinching mechanisms of templated carboxysome replication. For example, our pulse-chase data indicates that mature carboxysomes can fracture, as we observe repartitioning of RuBisCO to two daughter carboxysomes (Figure 6F).

While our data support some forms of replicative biogenesis, our methods cannot discriminate between *de novo* and templated nucleation events. It is plausible that RuBisCO seeds assembled *de novo* may be brought into close proximity with preexisting carboxysomes by other mechanisms. The carboxysome itself may be sufficient to capture independent seeds: crystal packing evidence from other studies suggests that shell proteins may contact one another face-to-face or assemble into antiparallel strips [11,24]. This may expose cargo-interacting surfaces to the outside of the carboxysome, creating affinity for cargo on the exterior as well as the interior carboxysome surface (Figure 6E).

In summary, all proposed mechanisms rely on the self-association of RuBisCO as the primary organizer and driving force of carboxysome biogenesis, with shell protein defining organelle geometry.

Broader implications

In addition to being crucial for global carbon fixation, the carboxysome has been proposed as a potential protein nano-factory capable of compartmentalizing heterologous reactions for metabolic engineering purposes [25]. An N-terminal peptide has been identified for the targeting of cargoes to 1,2-propanedeiol utilization microcompartments [26], but such a mechanism in carboxysomes has been elusive. Our studies of an assembly process dependent on self-association of cargo and the establishment of a unique internal microenvironment will inform the design of any future systems.

Materials and Methods

Bacterial strains and growth conditions

A table of all relevant strains and plasmids is presented in Table 1. All chemicals were obtained from Sigma-Aldrich unless otherwise noted (St. Louis, MO). The wild-type *Synechococcus elongatus* PCC 7942 strain was acquired from the American Type Culture Collection (ATCC, Manassas, VA). *S. elongatus* cells were grown in solid BG11 medium with an illumination of 2000 lux at 30°C [27]. *S. elongatus* were transformed following standard protocols by washing with 10mM sodium chloride followed by incubation overnight in the dark with 100 ng of plasmid DNA and subsequently plating on selective media [28]. Antibiotics were used at the following concentrations: kanamycin 10 µg/ml, spectinomycin 50 µg/ml, and chloramphenicol 10µg/ml. 25µM isopropyl β-D-1-thiogalactopyranoside (IPTG) induction was used for RbcL-GFP or RbcL-mOrange. 50µM IPTG was used to induce formation of bar carboxysomes and 1mM IPTG for RbcL-roGFP.

Plasmid construction

Cloning was done using Gibson assembly unless otherwise noted. IPTG inducible GFP strain (pDFS724) was obtained as previously described [16]. This neutral site 2 (NS2) plasmid contains a region with lacI and a promoter from pTRC99a followed by RbcL-sfGFP. sfGFP in pDFS724 was replaced by mOrange2 to obtain pAHC003. The shell protein fusion pAHC134 was obtained by modifying pDFS594s [16], replacing YFP with sfGFP. Two color strains were obtained by double transformation of pAHC003 and pAHC134. RbcL-SNAP fusion plasmid pAHC126 was obtained by replacing sfGFP in pDFS724 with SNAP tag. RbcL- roGFP fusion plasmid pAHC149 was constructed using restriction cloning at NheI and NotI. sfGFP in pDFS724 was replaced with roGFP1 (University of Oregon Remington Laboratory).

Image acquisition

Cells were plated onto BG11 + 2% agarose pads with IPTG as necessary and placed on a glass bottom dish (Part No.

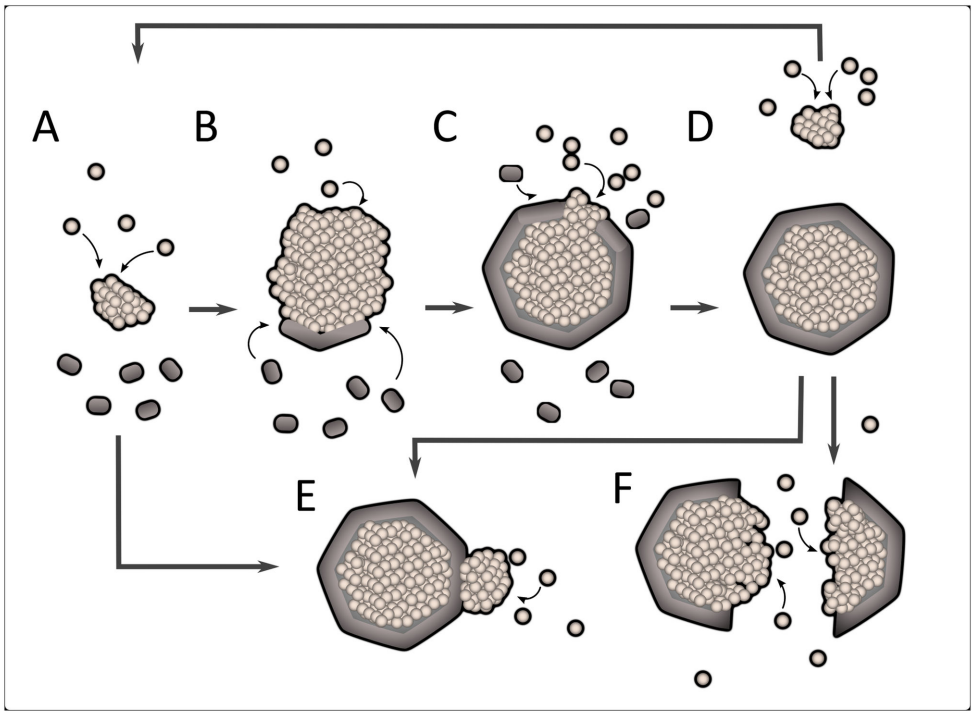

Figure 6. Model of carboxysome assembly. (A) RuBisCO seeds assemble from protomers over time. (B) Late in the assembly process, shell proteins rapidly assemble around RuBisCO. (C) Shell closure completes the carboxysome to establish an oxidizing environment, sealing RuBisCO from the cytosol. (D) A new RuBisCO nucleus forms after completion of the previous carboxysome. Colocalization may be driven by bisection of excess cargo by shell closure, or (E) by affinity of RuBisCO assemblies initiated elsewhere to the outside of the shell. (F) Rupture of a complete carboxysome would expose old RuBisCO cargo to template new assembly.

doi: 10.1371/journal.pone.0076127.g006

P35G-1.5-20-C, MatTek, Ashland, MA). The addition of 100µl of water around dish edges and a paraffin film seal permitted long-term imaging.

FRAP image acquisition was performed on a Nikon Ti inverted microscope (Nikon Instruments, Melville, NY) with a MicroPoint laser targeting system (Photonics Instruments, Saint Charles, IL) controlling a 100mW solid state 488nm laser for photobleaching. Imaging was performed using a 100x 1.4 numerical aperture objective, an EXFO XL-120 (Lumen Dynamics Group, Mississauga, Canada) fluorescence light source, and an ORCA-R2 charge coupled device camera (Hamamatsu Photonics, Hamamatsu, Japan).

As previously described [16], all other imaging was done using a Nikon TE-2000 microscope with a 100x 1.4 numerical aperture objective, a Lumencor LED fluorescence illuminator, and an ORCA-ER (Hamamatsu Photonics, Hamamatsu, Japan) charge coupled device camera. Acquisition was controlled using a custom MATLAB script controlling µManager [29] and a network AC power controller (IP Power 9258T) for photosynthetic lighting. Images were processed and analyzed with ImageJ.

Image analysis

For carboxysome lineage mapping, we used uTracker [30] to identify and localize closely-spaced point spread functions. The coordinates of these particles were imported into TrackMate, a plugin for FIJI [31] for manual annotation of track splitting events. Plots of lineages were retraced into vector format for counting division events and measuring maternal age.

For FRAP quantitation, the photobleaching rate after background subtraction was approximated with a linear function, which was used to correct measurements of the

Table 1. Bacterial Strains and Plasmids.

Strain or Plasmid	Relevant genotype	Resistance	Reference
E coli strains			
DH5-α	Host strain for plasmid construction		
S. elongatus strains			
PCC 7942	Wild-type *Synechococcus*, ATCC organism 33912		(Allen 1968)
RuBisCO/shell protein two color	Papca::ccmk4::sfGFP inserted in neutral site 1; lacI and ptrc::rbcL::mOrange2 inserted in neutral site 2	Kan/Sp	This work
Plasmids			
pDFS724	rbcL::sfGFP cloned into Neutral Site 2 at XbaI and NotI sites	Kan	(Savage 2010)
pDFS594S	Papca::ccmk4::YFP cloned into Neutral Site 1 at SpeI and NotI sites	Sp	(Savage 2010)
pAHC003	rbcL::mOrange2 cloned into pDFS724 in place of rbcL::sfGFP	Kan	This work
pAHC134	Papca::ccmk4::sfGFP	Sp	This work
pAHC126	rbcL::SNAP cloned into pDFS724 in place of rbcL::sfGFP	Kan	This work
pAHC149	rbcL::roGFP1 cloned into pDFS724 in place of rbcL::sfGFP	Kan	This work

bleached region. Intensity was normalized to the maximum (in frame 1).

For ratiometric imaging, we background subtracted both 410nm and 488nm images with a 50-pixel radius rolling ball. We then registered the images with translations measured from imaging fluorescent beads and divided the 488nm image by the 410nm after converting to 32 bit format for floating point operations. This images was then multiplied by a mask of carboxysomes we generated based on a thresholded, 10-pixel rolling ball radius background-subtracted 410nm image. The mean intensities of regions larger than 9 pixels^2 were quantitated with the "analyze particles" features of FIJI.

Pulse-chase SNAP dye labeling

RbcL-SNAP strains that were induced with 25μM IPTG for 12 to 24 hours were labeled with SNAP-Cell BG 505-Star (New England Biolabs, Ipswitch, MA) following manufacturer's instructions. Briefly, 1mL of cells were spun down and resuspended in 100 μL BG11 with 25μM IPTG and 5μM dye substrate. Labeling was done for 30 minutes in light. Cells were washed 3 times with BG11 and resuspended in BG11 with 25μM IPTG for 30 minutes in light. Cells were washed once more with BG11 and then transferred to an agarose pad for imaging. Either time-lapse imaging at 1 hour intervals or two endpoints 1 day apart were taken. Analysis was performed by manually counting and measuring foci intensity in ImageJ.

Western Blotting

Cells were lysed by sonication in 3% SDS lysis buffer, and proteins were separated on NuPAGE Novex 4-20% Tris-glycine gels (Life Technologies, Grand Island, NY). Transfer to a nitrocellulose membrane was performed using the iBlot Gel Transfer Device and iBlot Gel Transfer Stacks (Life technologies, Grand Island, NY). Subsequent blotting was done using the SNAP-ID Protein Detection System (EMD Millipore, Bellerica, MA) following manufacturer's instructions. Polyclonal anti-RuBisCO antibody (Agrisera, Prod. ID AS03 037) was used at a final dilution of 1:5000 and an HRP-conjugated goat anti-mouse antibody (Abcam, ab97265) was used at a final dilution of 1:5000. Peroxidase conjugates were detected using SuperSignal West Dura Exteded Duration substrate (Thermo Scientific).

Supporting Information

Figure S1. Quantification of RbcL and RbcL-GFP levels by Western blot. The inducible RbcL-GFP strain was grown in the presence or absence of 25μM IPTG at early log phase for 12 hours. Using a rabbit polyclonal anti-RuBisCO antibody, the intensities of bands above background were quantified; the RbcL-GFP band is 11% of the intensity of the endogenous RbcL band.
(JPG)

Figure S2. Mean and maximum velocities of carboxysome motion. 126 carboxysomes were tracked over a minimum of 85 frames, and the mean and maximum velocity of each track quantified. (A) The tracks overlaid on one frame of the movie. Different colors represent different tracks. (B) The mean of the mean carboxysome velocity is 11.5nm/minute. (C) The mean of the maximum carboxysome velocity is 60.6nm/minute. (D) Velocity is variable across each track, with a subset showing maximal velocity near the start of the track. (E) 25 tracks were selected at random from this set, and velocity was plotted against the frame number (ie, age) of each track. Interval of acquisition, 5 minute. (F) The mean velocity per frame number across these 25 tracks. As track length is variable, fewer data points contribute to the mean toward higher frame numbers.
(TIFF)

Figure S3. Bar carboxysomes colocalize with shell but are not oxidized. (A) Bar carboxysomes contain both RuBisCO and shell protein. Red, CcmK4-GFP. Green, RbcL-mOrange. Scale bar, 1μm. (B) Bar carboxysomes are relatively reduced compared to punctate carboxysomes. Still frame composite images of 488nm (reducing, green) and 408nm (oxidizing, purple) RbcL-roGFP1 as in Figure 4. Scale bar, 1μm.
(TIF)

Figure S4. Additional bar carboxysome FRAP data. Bleaching events are indicated by grey lines. Unbleached portions of the bar were used to correct for photobleaching.
(TIFF)

Figure S5. Additional shell protein assembly data. (A–C) Individual traces of the fluorescence intensity of CcmK4 foci. Each panel represents a different cell, and only shell foci in the

process of assembling are represented. Time interval: 5 minutes.
(TIF)

Video S1. Normal and bar carboxysomes are born from replicative events. Division events occur 2 seconds after the appearance of a white asterisk ~2μm above the relevant carboxysome. Imaging was initiated ~1 hour after induction. Green: RbcL-GFP. Red: phase contrast. Scale bar, 2μm. Frame rate, 12 frames (5 minute)/second.
(AVI)

Video S2. Shell protein is late to localize to RuBisCO assemblies. Shell localization events occur 2 seconds after the appearance of a white asterisk ~1μm above the relevant carboxysome. The top-most highlighted carboxysome also nucleates a daughter at the 12 hour mark. Imaging was initiated ~3 hours after induction. Green: RbcL-mOrange. Red: CcmK4-GFP. Scale bar, 2μm. Frame rate, 7 frames (5 minute)/second.
(AVI)

Video S3. Nascent, reduced carboxysomes oxidize as they mature. Oxidation events occur 2 seconds after the appearance of a white asterisk ~1μm above the relevant carboxysome. Imaging was initiated ~24 hours after induction. Green: 488Ex RbcL-roGFP1. Magenta: 410Ex RbcL-roGFP1. Scale bar, 2μm. Frame rate, 7 frames (10 minute)/second.
(AVI)

Acknowledgements

We thank Jennifer Waters and Josh Rosenberg (Harvard) for advice and technical assistance with FRAP and ratiometric imaging at the Nikon Imaging Center at Harvard Medical School, Burak Orkmus (Harvard) for the generous loan of optical components, and Jeff Way, Tim Mitchison, Tom Shemesh, Jacob Stewart-Ornstein, and Joe Torella (Harvard) for many helpful discussions. Finally, we are grateful to Ethan Garner (Harvard) for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: AHC DFS PAS JKP. Performed the experiments: AHC JKP. Analyzed the data: AHC ARM JKP. Contributed reagents/materials/analysis tools: AHC ARM DHS. Wrote the manuscript: AHC DHS PAS JKP.

References

1. Fuerst JA, Webb RI (1991) Membrane-bounded nucleoid in the Eubacterium Gemmata obscuriglobus. Proc Natl Acad Sci U S A 88: 8184–8188. doi:10.1073/pnas.88.18.8184. PubMed: 11607213.
2. Mullineaux CW (1999) The thylakoid membranes of cyanobacteria: structure, dynamics and function. Funct Plant Biol 26: 671–677.
3. Murat D, Quinlan A, Vali H, Komeili A (2010) Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. Proc Natl Acad Sci U S A 107: 5593–5598. doi:10.1073/pnas.0914439107. PubMed: 20212111.
4. Pfeifer F (2012) Distribution, formation and regulation of gas vesicles. Nat Rev Microbiol 10: 705–715. doi:10.1038/nrmicro2834. PubMed: 22941504.
5. Kerfeld CA, Heinhorst S, Cannon GC (2010) Bacterial Microcompartments. Annu Rev Microbiol 64: 391–408. doi:10.1146/annurev.micro.112408.134211. PubMed: 20825353.
6. Yeates TO, Crowley CS, Tanaka S (2010) Bacterial microcompartment organelles: protein shell structure and evolution. Annu Rev Biophys 39: 185–205. doi:10.1146/annurev.biophys.093008.131418. PubMed: 20192762.
7. Shively JM, Ball F, Brown DH, Saunders RE (1973) Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. Science 182: 584–586. doi:10.1126/science.182.4112.584. PubMed: 4355679.
8. Codd GA, Marsden WJN (1984) The Carboxysomes (polyhedral Bodies) of Autotrophic Prokarygtes. Biol Rev 59: 389–422. doi:10.1111/j.1469-185X.1984.tb00710.x.
9. Cannon GC, Bradburne CE, Aldrich HC, Baker SH, Heinhorst S et al. (2001) Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra. Appl Environ Microbiol 67: 5351–5361. doi:10.1128/AEM.67.12.5351-5361.2001. PubMed: 11722879.
10. Peña KL, Castel SE, Araujo C de, Espie GS, Kimber MS (2010) Structural basis of the oxidative activation of the carboxysomal γ-carbonic anhydrase, CcmM. Proc Natl Acad Sci USA 107: 2455–2460. doi:10.1073/pnas.0910866107. PubMed: 20133749.
11. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S et al. (2008) Atomic-level models of the bacterial carboxysome shell. Science 319: 1083–1086. doi:10.1126/science.1151458. PubMed: 18292340.
12. Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M et al. (2005) Protein Structures Forming the Shell of Primitive Bacterial Organelles. Science 309: 936–938. doi:10.1126/science.1113397. PubMed: 16081736.
13. Iancu CV, Morris DM, Dou Z, Heinhorst S, Cannon GC et al. (2010) Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. J Mol Biol 396: 105–117. doi:10.1016/j.jmb.2009.11.019. PubMed: 19925807.
14. Orus MI, Rodriguez ML, Martinez F, Marco E (1995) Biogenesis and Ultrastructure of Carboxysomes from Wild Type and Mutants of Synechococcus sp. Strain PCC 7942. Plant Physiol 107: 1159–1166.
15. Menon BB, Dou Z, Heinhorst S, Shively JM, Cannon GC (2008) Halothiobacillus neapolitanus carboxysomes sequester heterologous and chimeric RubisCO species. PLOS ONE 3: e3570. doi:10.1371/journal.pone.0003570. PubMed: 18974784.
16. Savage DF, Afonso B, Chen AH, Silver PA (2010) Spatially Ordered Dynamics of the Bacterial Carbon Fixation Machinery. Science 327: 1258–1261. doi:10.1126/science.1186090. PubMed: 20203050.
17. Gantt E, Conti SF (1969) Ultrastructure of Blue-Green Algae1. J Bacteriol 97: 1486–1493. PubMed: 5776533.
18. McKay RML, Gibbs SP, Espie GS (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.
19. Giglia-Mari G, Theil AF, Mari P-O, Mourgues S, Nonnekens J et al. (2009) Differentiation Driven Changes in the Dynamic Organization of Basal Transcription Initiation. PLOS Biol 7: e1000220. doi:10.1371/journal.pbio.1000220. PubMed: 19841728.
20. Peña KL, Castel SE, de Araujo C, Espie GS, Kimber MS (2010) Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM. Proc Natl Acad Sci U S A 107: 2455–2460. doi:10.1073/pnas.0910866107. PubMed: 20133749.
21. Hanson GT, Aggeler R, Oglesbee D, Cannon M, Capaldi RA et al. (2004) Investigating Mitochondrial Redox Potential with Redox-sensitive Green Fluorescent Protein Indicators. J Biol Chem 279: 13044–13053. doi:10.1074/jbc.M312846200. PubMed: 14722062.
22. Iancu CV, Ding HJ, Morris DM, Dias DP, Gonzales AD et al. (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. PubMed: 17669419.
23. Schmid MF, Paredes AM, Khant HA, Soyer F, Aldrich HC et al. (2006) Structure of Halothiobacillus neapolitanus Carboxysomes by Cryo-electron Tomography. J Mol Biol 364: 526–535. doi:10.1016/j.jmb.2006.09.024. PubMed: 17028023.

24. Cai F, Sutter M, Cameron JC, Stanley DN, Kinney JN et al. (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. PubMed: 23572529.
25. Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P et al. (2012) Modularity of a carbon-fixing protein organelle. Proc Natl Acad Sci U S A 109: 478–483. doi:10.1073/pnas.1108557109. PubMed: 22184212.
26. Fan C, Cheng S, Liu Y, Escobar CM, Crowley CS et al. (2010) Short N-terminal sequences package proteins into bacterial microcompartments. Proc Natl Acad Sci USA 107: 7509–7514. doi:10.1073/pnas.0913199107. PubMed: 20308536.
27. Allen MM, Stanier RY (1968) Growth and Division of Some Unicellular Blue-green Algae. J Gen Microbiol 51: 199–202. doi:10.1099/00221287-51-2-199. PubMed: 5652095.
28. Clerico EM, Ditty JL, Golden SS (2007) Specialized techniques for site-directed mutagenesis in cyanobacteria. Methods Mol Biol Clifton NJ 362: 155–171. doi:10.1007/978-1-59745-257-1_11. PubMed: 17417008.
29. Edelstein A, Amodaj N, Hoover K, Vale R, Stuurman N (2001) Computer Control of Microscopes Using μManager. Current Protocols in Molecular Biology. John Wiley & Sons, Inc. Available: http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1002/0471142727.mb1420s92/abstract. Accessed 11 July 2013.
30. Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S et al. (2008) Robust single particle tracking in live cell time-lapse sequences. Nat Methods 5: 695–702. doi:10.1038/nmeth.1237. PubMed: 18641657.
31. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M et al. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9: 676–682. doi:10.1038/nmeth.2019. PubMed: 22743772.

References

[1] (2007). Spatiotemporal mechanisms of life. *Nature chemical biology*, 3(10), 593.

[2] Allen, M. M. & Stanier, R. Y. (1968). Growth and Division of Some Unicellular Blue-green Algae. *Journal of General Microbiology*, 51(2), 199–202.

[3] Badger, M. R. (2003). CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *Journal of Experimental Botany*, 54(383), 609–622.

[4] Barber, J. (2009). Photosynthetic energy conversion: natural and artificial. *Chemical Society reviews*, 38(1), 185–96.

[5] Bates, D. (2008). The bacterial replisome: back on track? *Molecular microbiology*, 69(6), 1341–8.

[6] Bayer, T. S., Widmaier, D. M., Temme, K., Mirsky, E. A., Santi, D. V., & Voigt, C. A. (2009). Synthesis of methyl halides from biomass using engineered microbes. *Journal of the American Chemical Society*, 131(18), 6508–15.

[7] Berlatzky, I. A., Rouvinski, A., & Ben-Yehuda, S. (2008). Spatial organization of a replicating bacterial chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37), 14136–40.

[8] Binder, B. J. & Chisholm, S. W. (1990). Relationship between DNA cycle and growth rate in Synechococcus sp. strain PCC 6301. *J. Bacteriol.*, 172(5), 2313–2319.

[9] Bonacci, W., Teng, P. K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P. A., & Savage, D. F. (2012). Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences of the United States of America*, 109(2), 478–83.

[10] Brophy, J. A. N. & Voigt, C. A. (2014). Principles of genetic circuit design. *Nature Methods*, 11(5), 508–520.

[11] Cai, F., Sutter, M., Cameron, J. C., Stanley, D. N., Kinney, J. N., & Kerfeld, C. A. (2013). The structure of CcmP, a tandem bacterial microcompartment domain protein from the β-carboxysome, forms a subcompartment within a microcompartment. *The Journal of biological chemistry*, 288(22), 16055–63.

[12] Cannon, G. C., Bradburne, C. E., Aldrich, H. C., Baker, S. H., Heinhorst, S., & Shively, J. M. (2001). Microcompartments in prokaryotes: carboxysomes and related polyhedra. *Applied and environmental microbiology*, 67(12), 5351–61.

[13] Cannon, G. C. & Shively, J. M. (1983). Characterization of a homogenous preparation of carboxysomes from Thiobacillus neapolitanus. *Archives of Microbiology*, 134(1), 52–59.

[14] Chang, Y.-F. & Chang, C.-H. (2011). CAGO: a software tool for dynamic visual comparison and correlation measurement of genome organization. *PloS one*, 6(11), e27080.

[15] Chang Z., S. P. A. C. Y. Y. (2015). Identification and selective expansion of functionally superior T cells expressing chimeric antigen receptors.

[16] Chapman, A. G., Fall, L., & Atkinson, D. E. (1971). Adenylate Energy Charge in Escherichia coli During Growth and Starvation. *J. Bacteriol.*, 108(3), 1072–1086.

[17] Chen, A. H., Afonso, B., Silver, P. A., & Savage, D. F. (2012). Spatial and temporal organization of chromosome duplication and segregation in the cyanobacterium Synechococcus elongatus PCC 7942. *PloS one*, 7(10), e47837.

[18] Chen, H.-N. & Woycechowsky, K. J. (2012). Conversion of a Dodecahedral Protein Capsid into Pentamers via Minimal Point Mutations. *Biochemistry*, (pp. 120518124810004).

[19] Chen, Y. Y., Jensen, M. C., & Smolke, C. D. (2010). Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19), 8531–6.

[20] Choudhary, S., Quin, M. B., Sanders, M. A., Johnson, E. T., & Schmidt-Dannert, C. (2012). Engineered Protein Nano-Compartments for Targeted Enzyme Localization. *PLoS ONE*, 7(3), e33342.

[21] Clerico, E. M., Ditty, J. L., & Golden, S. S. (2007). Specialized techniques for site-directed mutagenesis in cyanobacteria. *Methods in molecular biology (Clifton, N.J.)*, 362, 155–71.

[22] CODD, G. A. & MARSDEN, W. J. N. (1984). THE CARBOXYSOMES (POLYHEDRAL BODIES) OF AUTOTROPHIC PROKARYGTES. *Biological Reviews*, 59(3), 389–422.

[23] Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339(6121), 819–23.

[24] Conrado, R. J., Varner, J. D., & DeLisa, M. P. (2008). Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current opinion in biotechnology*, 19(5), 492–499.

[25] Conrado, R. J., Wu, G. C., Boock, J. T., Xu, H., Chen, S. Y., Lebar, T., Turnsek, J., Tomsic, N., Avbelj, M., Gaber, R., Koprivnjak, T., Mori, J., Glavnik, V., Vovk, I., Bencina, M., Hodnik, V., Anderluh, G., Dueber, J. E., Jerala, R., & Delisa, M. P. (2012). DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic acids research*, 40(4), 1879–1889.

[26] Damiola, F. (2000). Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes & Development*, 14(23), 2950–2961.

[27] Delebecque, C. J., Lindner, A. B., Silver, P. A., & Aldaye, F. A. (2011). Organization of intracellular reactions with rationally designed RNA assemblies. *Science (New York, N.Y.)*, 333(6041), 470–4.

[28] Dibner, C., Schibler, U., & Albrecht, U. (2010). The Mammalian Circadian Timing System: Organization and Coordination of Central and Peripheral Clocks. *Annual Review of Physiology*, 72(1), 517–549.

[29] Dove, S. L. & Hochschild, A. (2004). A bacterial two-hybrid system based on transcription activation. *Methods in molecular biology (Clifton, N.J.)*, 261, 231–46.

[30] Dove, S. L., Joung, J. K., & Hochschild, A. (1997). Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature*, 386(6625), 627–30.

[31] Ducat, D. C., Avelar-Rivas, J. A., Way, J. C., & Silver, P. A. (2012). Rerouting carbon flux to enhance photosynthetic productivity. *Applied and environmental microbiology*, 78(8), 2660–8.

[32] Ducat, D. C., Way, J. C., & Silver, P. A. (2011). Engineering cyanobacteria to generate high-value products. *Trends in biotechnology*, 29(2), 95–103.

[33] Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., Prather, K. L. J., & Keasling, J. D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nature biotechnology*, 27(8), 753–9.

[34] Dutta, D., De, D., Chaudhuri, S., & Bhattacharya, S. K. (2005). Hydrogen production by Cyanobacteria. *Microbial cell factories*, 4(1), 36.

[35] Ebersbach, G. & Gerdes, K. (2005). Plasmid segregation mechanisms. *Annual review of genetics*, 39, 453–79.

[36] Edelstein, A. D., Tsuchida, M. A., Amodaj, N., Pinkard, H., Vale, R. D., & Stuurman, N. (2014). Advanced methods of microscope control using μManager software. *Journal of Biological Methods*, 1(2), 10.

[37] Edgar, R. S., Green, E. W., Zhao, Y., van Ooijen, G., Olmedo, M., Qin, X., Xu, Y., Pan, M., Valekunja, U. K., Feeney, K. A., Maywood, E. S., Hastings, M. H., Baliga, N. S., Merrow, M., Millar, A. J., Johnson, C. H., Kyriacou, C. P., O'Neill, J. S., & Reddy, A. B. (2012). Peroxiredoxins are conserved markers of circadian rhythms. *Nature*, 485(7399), 459–64.

[38] Eichelmann, H., Talts, E., Oja, V., Padu, E., & Laisk, A. (2009). Rubisco in planta kcat is regulated in balance with photosynthetic electron transport. *Journal of experimental botany*, 60(14), 4077–88.

[39] Eisenhut, M., Kahlon, S., Hasse, D., Ewald, R., Lieman-Hurwitz, J., Ogawa, T., Ruth, W., Bauwe, H., Kaplan, A., & Hagemann, M. (2006). The plant-like C2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria. *Plant physiology*, 142(1), 333–42.

[40] Ellis, R. (1979). The most abundant protein in the world. *Trends in Biochemical Sciences*, 4(11), 241–244.

[41] Elowitz, M. B. & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), 335–8.

[42] Eriksson, M. & Clarke, A. (1996). The heat shock protein ClpB mediates the development of thermotolerance in the cyanobacterium Synechococcus sp. strain PCC 7942. *J. Bacteriol.*, 178(16), 4839–4846.

[43] Falkowski, P. G. & Raven, J. A. (2013). *Aquatic Photosynthesis: (Second Edition)*. Princeton University Press.

[44] Fan, C., Cheng, S., Liu, Y., Escobar, C. M., Crowley, C. S., Jefferson, R. E., Yeates, T. O., & Bobik, T. A. (2010). Short N-terminal sequences package proteins into bacterial microcompartments. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16), 7509–14.

[45] Farhi, M., Marhevka, E., Masci, T., Marcos, E., Eyal, Y., Ovadis, M., Abeliovich, H., & Vainstein, A. (2011). Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metabolic engineering*, 13(5), 474–81.

[46] Fogel, M. A. & Waldor, M. K. (2005). Distinct segregation dynamics of the two Vibrio cholerae chromosomes. *Molecular microbiology*, 55(1), 125–36.

[47] Fogel, M. A. & Waldor, M. K. (2006). A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes & development*, 20(23), 3269–82.

[48] Fu, J., Yang, Y. R., Johnson-Buck, A., Liu, M., Liu, Y., Walter, N. G., Woodbury, N. W., & Yan, H. (2014). Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nature nanotechnology*, 9(7), 531–6.

[49] Fuerst, J. A. & Webb, R. I. (1991). Membrane-bounded nucleoid in the eubacterium Gemmata obscuriglobus. *Proceedings of the National Academy of Sciences of the United States of America*, 88(18), 8184–8.

[50] Gachon, F., Nagoshi, E., Brown, S. a., Ripperger, J., & Schibler, U. (2004). The mammalian circadian timing system: from gene expression to physiology. *Chromosoma*, 113(3), 103–112.

[51] Gantt, E. & Conti, S. F. (1969). Ultrastructure of blue-green algae. *Journal of bacteriology*, 97(3), 1486–93.

[52] Gao, F. & Zhang, C.-T. (2008). Ori-Finder: a web-based system for finding oriCs in unannotated bacterial genomes. *BMC bioinformatics*, 9(1), 79.

[53] Gehrmann, W. & Elsner, M. (2011). A specific fluorescence probe for hydrogen peroxide detection in peroxisomes. *Free Radical Research*, 45(5), 6.

[54] Gerdes, K., Howard, M., & Szardenings, F. (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell*, 141(6), 927–42.

[55] Germain, A. & Kupfer, D. J. (2008). Circadian rhythm disturbances in depression. *Human psychopharmacology*, 23(7), 571–85.

[56] Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, 6(5), 343–5.

[57] Giglia-Mari, G., Theil, A. F., Mari, P.-O., Mourgues, S., Nonnekens, J., Andrieux, L. O., de Wit, J., Miquel, C., Wijgers, N., Maas, A., Fousteri, M., Hoeijmakers, J. H. J., & Vermeulen, W. (2009). Differentiation Driven Changes in the Dynamic Organization of Basal Transcription Initiation. *PLoS Biology*, 7(10), e1000220.

[58] Gitai, Z., Thanbichler, M., & Shapiro, L. (2005). The choreographed dynamics of bacterial chromosomes. *Trends in microbiology*, 13(5), 221–8.

[59] Gordon, G. S., Shivers, R. P., & Wright, A. (2002). Polar localization of the Escherichia coli oriC region is independent of the site of replication initiation. *Molecular microbiology*, 44(2), 501–7.

[60] Graumann, P. L., Dame, R. T., & Dorman, C. J. (2010). The Chromosome Segregation Machinery in Bacteria. In *Bacterial Chromatin (Google eBook)* chapter 3, (pp. 31–48). Springer.

[61] Griese, M., Lange, C., & Soppa, J. (2011). Ploidy in cyanobacteria. *FEMS Microbiology Letters*, 323(2), 124–131.

[62] Gutu, A. & O'Shea, E. K. (2013). Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. *Molecular cell*, 50(2), 288–94.

[63] Ha, T., Kozlov, A. G., & Lohman, T. M. (2012). Single-Molecule Views of Protein Movement on Single-Stranded DNA.

[64] Haeusser, D. P. & Levin, P. A. (2008). The great divide: coordinating cell cycle events during bacterial growth and division. *Current opinion in microbiology*, 11(2), 94–9.

[65] Hanaoka, M., Takai, N., Hosokawa, N., Fujiwara, M., Akimoto, Y., Kobori, N., Iwasaki, H., Kondo, T., & Tanaka, K. (2012). RpaB, another response regulator operating circadian clock-dependent transcriptional regulation in Synechococcus elongatus PCC 7942. *The Journal of biological chemistry*, 287(31), 26321–7.

[66] Hanson, G. T., Aggeler, R., Oglesbee, D., Cannon, M., Capaldi, R. A., Tsien, R. Y., & Remington, S. J. (2004). Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *The Journal of biological chemistry*, 279(13), 13044–53.

[67] Heidorn, T., Camsund, D., Huang, H.-H., Lindberg, P., Oliveira, P., Stensjö, K., & Lindblad, P. (2011). Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods in enzymology*, 497, 539–79.

[68] Heinhorst, S., Williams, E. B., Cai, F., Murin, C. D., Shively, J. M., & Cannon, G. C. (2006). Characterization of the carboxysomal carbonic anhydrase CsoSCA from Halothiobacillus neapolitanus. *Journal of bacteriology*, 188(23), 8087–94.

[69] Holtman, C. K. (2005). High-Throughput Functional Analysis of the Synechococcus elongatus PCC 7942 Genome. *DNA Research*, 12(2), 103–115.

[70] Hu, B., Yang, G., Zhao, W., Zhang, Y., & Zhao, J. (2007). MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium Anabaena sp. PCC 7120. *Molecular microbiology*, 63(6), 1640–52.

[71] Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., & Davies, D. R. (1988). Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *J. Biol. Chem.*, 263(33), 17857–17871.

[72] Iancu, C. V., Ding, H. J., Morris, D. M., Dias, D. P., Gonzales, A. D., Martino, A., & Jensen, G. J. (2007). The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. *Journal of molecular biology*, 372(3), 764–73.

[73] Iancu, C. V., Morris, D. M., Dou, Z., Heinhorst, S., Cannon, G. C., & Jensen, G. J. (2010). Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. *Journal of molecular biology*, 396(1), 105–17.

[74] Ishiura, M. (1998). Expression of a Gene Cluster kaiABC as a Circadian Feedback Process in Cyanobacteria. *Science*, 281(5382), 1519–1523.

[75] Ito, H., Kageyama, H., Mutsuda, M., Nakajima, M., Oyama, T., & Kondo, T. (2007). Autonomous synchronization of the circadian KaiC phosphorylation rhythm. *Nature structural & molecular biology*, 14(11), 1084–8.

[76] Iwasaki, H., Nishiwaki, T., Kitayama, Y., Nakajima, M., & Kondo, T. (2002). KaiA-stimulated KaiC phosphorylation in circadian timing loops in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24), 15788–93.

[77] Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S. L., & Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods*, 5(8), 695–702.

[78] Jensen, R. B. (2006). Coordination between chromosome replication, segregation, and cell division in Caulobacter crescentus. *Journal of bacteriology*, 188(6), 2244–53.

[79] Jensen, R. B. & Shapiro, L. (1999). The Caulobacter crescentus smc gene is required for cell cycle progression and chromosome segregation. *Proceedings of the National Academy of Sciences*, 96(19), 10661–10666.

[80] Jensen, R. B., Wang, S. C., & Shapiro, L. (2001). A moving DNA replication factory in Caulobacter crescentus. *The EMBO journal*, 20(17), 4952–63.

[81] Jø rgensen, K., Rasmussen, A. V., Morant, M., Nielsen, A. H., Bjarnholt, N., Zagrobelny, M., Bak, S. r., & Mø ller, B. L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current opinion in plant biology*, 8(3), 280–91.

[82] Kaasik, K. & Chi Lee, C. (2004). Reciprocal regulation of haem biosynthesis and the circadian clock in mammals. *Nature*, 430(6998), 467–471.

[83] Kerfeld, C. A., Heinhorst, S., & Cannon, G. C. (2010). Bacterial microcompartments. *Annual review of microbiology*, 64, 391–408.

[84] Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M., & Yeates, T. O. (2005). Protein structures forming the shell of primitive bacterial organelles. *Science (New York, N.Y.)*, 309(5736), 936–8.

[85] Kim, Y.-I., Dong, G., Carruthers, C. W., Golden, S. S., & LiWang, A. (2008). The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 105(35), 12825–30.

[86] Kinney, J. N., Salmeen, A., Cai, F., & Kerfeld, C. A. (2012). Elucidating the essential role of the conserved carboxysomal protein CcmN reveals a common feature of bacterial microcompartment assembly. *The Journal of biological chemistry*.

[87] Kirschner, M., Newport, J., & Gerhart, J. (1985). The timing of early developmental events in Xenopus. *Trends in Genetics*, 1, 41–47.

[88] Kizer, L., Pitera, D. J., Pfleger, B. F., & Keasling, J. D. (2008). Application of functional genomics to pathway optimization for increased isoprenoid production. *Applied and environmental microbiology*, 74(10), 3229–41.

[89] Klein, D. C., Moore, R. Y., & Reppert, S. M. (1991). *Suprachiasmatic Nucleus: The Mind's Clock*. Oxford University Press.

[90] Klein, M. G., Zwart, P., Bagby, S. C., Cai, F., Chisholm, S. W., Heinhorst, S., Cannon, G. C., & Kerfeld, C. A. (2009). Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *Journal of molecular biology*, 392(2), 319–33.

[91] Kondo, T., Strayer, C. A., Kulkarni, R. D., Taylor, W., Ishiura, M., Golden, S. S., & Johnson, C. H. (1993). Circadian rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. *Proceedings of the National Academy of Sciences*, 90(12), 5672–5676.

[92] Lau, I. F., Filipe, S. R., Sø balle, B., Ø kstad, O.-A., Barre, F.-X., & Sherratt, D. J. (2003). Spatial and temporal organization of replicating Escherichia coli chromosomes. *Molecular microbiology*, 49(3), 731–43.

[93] Lau, I. F., Filipe, S. R., Sø balle, B., Ø kstad, O.-A., Barre, F.-X., & Sherratt, D. J. (2004). Spatial and temporal organization of replicating Escherichia coli chromosomes. *Molecular Microbiology*, 49(3), 731–743.

[94] Lévi, F. (2002). From circadian rhythms to cancer chronotherapeutics. *Chronobiology international*, 19(1), 1–19.

[95] Levi, F. & Schibler, U. (2007). Circadian rhythms: mechanisms and therapeutic implications. *Annual review of pharmacology and toxicology*, 47, 593–628.

[96] Levine, J., Funes, P., Dowse, H., & Hall, J. (2002). Signal analysis of behavioral and molecular cycles. *BMC Neuroscience*, 3(1), 1.

[97] Li, Y., Sergueev, K., & Austin, S. (2002). The segregation of the Escherichia coli origin and terminus of replication. *Molecular microbiology*, 46(4), 985–96.

[98] Lindsay, M., Webb, R., Strous, M., Jetten, M., Butler, M., Forde, R., & Fuerst, J. (2001). Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. *Archives of Microbiology*, 175(6), 413–429.

[99] Liu, Y., Beer, L. L., & Whitman, W. B. (2012). Methanogens: a window into ancient sulfur metabolism. *Trends in microbiology*.

[100] Liu, Y., Tsinoremas, N. F., Johnson, C. H., Lebedeva, N. V., Golden, S. S., Ishiura, M., & Kondo, T. (1995). Circadian orchestration of gene expression in cyanobacteria. *Genes & development*, 9(12), 1469–78.

[101] Lø bner Olesen, A., Skarstad, K., Hansen, F. G., von Meyenburg, K., & Boye, E. (1989). The DnaA protein determines the initiation mass of Escherichia coli K-12. *Cell*, 57(5), 881–889.

[102] Long, B. M., Badger, M. R., Whitney, S. M., & Price, G. D. (2007). Analysis of carboxysomes from Synechococcus PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *The Journal of Biological Chemistry*, 282(40), 29323–29335.

[103] Losick, R. & Shapiro, L. (1999). Changing Views on the Nature of the Bacterial Cell: from Biochemistry to Cytology. *J. Bacteriol.*, 181(14), 4143–4145.

[104] Machado, I. M. P. & Atsumi, S. (2012). Cyanobacterial biofuel production. *Journal of biotechnology*, 162(1), 50–6.

[105] Mackey, S. R., Ditty, J. L., Clerico, E. M., & Golden, S. S. (2007). Detection of rhythmic bioluminescence from luciferase reporters in cyanobacteria. *Methods in molecular biology (Clifton, N.J.)*, 362, 115–29.

[106] Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, 339(6121), 823–6.

[107] Mangan, N. & Brenner, M. (2014). Systems analysis of the CO2 concentrating mechanism in cyanobacteria. *eLife*, (pp. e02043).

[108] Manjasetty, B. A., Powlowski, J., & Vrielink, A. (2003). Crystal structure of a bifunctional aldolase-dehydrogenase: sequestering a reactive and volatile intermediate. *Proceedings of the National Academy of Sciences of the United States of America*, 100(12), 6992–7.

[109] Markson, J. S., Piechura, J. R., Puszynska, A. M., & O'Shea, E. K. (2013). Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell*, 155(6), 1396–408.

[110] Martin, W. (2010). Evolutionary origins of metabolic compartmentalization in eukaryotes. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1541), 847–55.

[111] McKay, R. M. L., Gibbs, S. P., & 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. *Archives of Microbiology*, 159(1), 21–29.

[112] Menon, B. B., Dou, Z., Heinhorst, S., Shively, J. M., & Cannon, G. C. (2008). Halothiobacillus neapolitanus carboxysomes sequester heterologous and chimeric RubisCO species. *PloS one*, 3(10), e3570.

[113] Mori, T., Binder, B., & Johnson, C. H. (1996). Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19), 10183–8.

[114] Mott, M. L. & Berger, J. M. (2007). DNA replication initiation: mechanisms and regulation in bacteria. *Nature reviews. Microbiology*, 5(5), 343–54.

[115] Mullineaux, C. W. (1999). The thylakoid membranes of cyanobacteria: structure, dynamics and function. *Australian Journal of Plant Physiology*, 26(7), 671.

[116] Murat, D., Quinlan, A., Vali, H., & Komeili, A. (2010). Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. *Proceedings of the National Academy of Sciences of the United States of America*, 107(12), 5593–8.

[117] Murray, A. (1994). Cell cycle checkpoints. *Current Opinion in Cell Biology*, 6(6), 872–876.

[118] Murray, D. B., Roller, S., Kuriyama, H., & Lloyd, D. (2001). Clock Control of Ultradian Respiratory Oscillation Found during Yeast Continuous Culture. *Journal of Bacteriology*, 183(24), 7253–7259.

[119] Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., & Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science (New York, N.Y.)*, 308(5720), 414–5.

[120] Niki, H. & Hiraga, S. (1998). Polar localization of the replication origin and terminus in Escherichia coli nucleoids during chromosome partitioning. *Genes & development*, 12(7), 1036–45.

[121] Niki, H., Yamaichi, Y., & Hiraga, S. (2000). Dynamic organization of chromosomal DNA in Escherichia coli. *Genes & Dev.*, 14(2), 212–223.

[122] Noirot-Gros, M.-F., Dervyn, E., Wu, L. J., Mervelet, P., Errington, J., Ehrlich, S. D., & Noirot, P. (2002). An expanded view of bacterial DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), 8342–7.

[123] Norman, T. M., Lord, N. D., Paulsson, J., & Losick, R. (2013). Memory and modularity in cell-fate decision making. *Nature*, 503(7477), 481–6.

[124] Norsker, N.-H., Barbosa, M. J., Vermuë, M. H., & Wijffels, R. H. (2011). Microalgal production–a close look at the economics. *Biotechnology advances*, 29(1), 24–7.

[125] Orus, M. I., Rodriguez, M. L., Martinez, F., & Marco, E. (1995). Biogenesis and Ultrastructure of Carboxysomes from Wild Type and Mutants of Synechococcus sp. Strain PCC 7942. *Plant physiology*, 107(4), 1159–1166.

[126] Park, D. H. (1999). Control of Circadian Rhythms and Photoperiodic Flowering by the Arabidopsis GIGANTEA Gene. *Science*, 285(5433), 1579–1582.

[127] Parsons, J. B., Frank, S., Bhella, D., Liang, M., Prentice, M. B., Mulvihill, D. P., & Warren, M. J. (2010). Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Molecular cell*, 38(2), 305–15.

[128] Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., & Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology*, 24(1), 79–88.

[129] Pena, K. L., Castel, S. E., de Araujo, C., Espie, G. S., & Kimber, M. S. (2010). Structural basis of the oxidative activation of the carboxysomal -carbonic anhydrase, CcmM. *Proceedings of the National Academy of Sciences*, 107(6), 2455–2460.

[130] Penrod, J. T. & Roth, J. R. (2006). Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. *Journal of bacteriology*, 188(8), 2865–74.

[131] Pfeifer, B. A. & Khosla, C. (2001). Biosynthesis of polyketides in heterologous hosts. *Microbiology and molecular biology reviews : MMBR*, 65(1), 106–18.

[132] Pfeifer, F. (2012). Distribution, formation and regulation of gas vesicles. *Nature reviews. Microbiology*, 10(10), 705–15.

[133] Phillips, I. & Silver, P. (2006). A New Biobrick Assembly Strategy Designed for Facile Protein Engineering.

[134] Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, 152(5), 1173–83.

[135] Rae, B. D., Long, B. M., Badger, M. R., & Price, G. D. (2013). Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria. *Microbiology and molecular biology reviews : MMBR*, 77(3), 357–79.

[136] Rascher, U., Hütt, M. T., Siebke, K., Osmond, B., Beck, F., & Lüttge, U. (2001). Spatiotemporal variation of metabolism in a plant circadian rhythm: the biological clock as an assembly of coupled individual oscillators. *Proceedings of the National Academy of Sciences of the United States of America*, 98(20), 11801–5.

[137] Reyes-Lamothe, R., Possoz, C., Danilova, O., & Sherratt, D. J. (2008a). Independent positioning and action of Escherichia coli replisomes in live cells. *Cell*, 133(1), 90–102.

[138] Reyes-Lamothe, R., Wang, X., & Sherratt, D. (2008b). Escherichia coli and its chromosome. *Trends in microbiology*, 16(5), 238–45.

[139] Ringgaard, S., van Zon, J., Howard, M., & Gerdes, K. (2009). Movement and equipositioning of plasmids by ParA filament disassembly. *Proceedings of the National Academy of Sciences of the United States of America*, 106(46), 19369–74.

[140] Roenneberg, T. & Merrow, M. (1998). Molecular circadian oscillators: an alternative hypothesis. *Journal of biological rhythms*, 13(2), 167–79.

[141] Rust, M. J., Golden, S. S., & O'Shea, E. K. (2011). Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science (New York, N.Y.)*, 331(6014), 220–3.

[142] Sampson, E. M. & Bobik, T. A. (2008). Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *Journal of bacteriology*, 190(8), 2966–71.

[143] Savage, D. F., Afonso, B., Chen, A. H., & Silver, P. A. (2010). Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science (New York, N.Y.)*, 327(5970), 1258–61.

[144] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7), 676–82.

[145] Schmid, M. F., Paredes, A. M., Khant, H. A., Soyer, F., Aldrich, H. C., Chiu, W., & Shively, J. M. (2006). Structure of Halothiobacillus neapolitanus carboxysomes by cryo-electron tomography. *Journal of molecular biology*, 364(3), 526–35.

[146] Schneider, D., Fuhrmann, E., Scholz, I., Hess, W. R., & Graumann, P. L. (2007). Fluorescence staining of live cyanobacterial cells suggest non-stringent chromosome segregation and absence of a connection between cytoplasmic and thylakoid membranes. *BMC cell biology*, 8(1), 39.

[147] Sehgal, A. (2004). *Molecular Biology of Circadian Rhythms*. John Wiley & Sons.

[148] Seki, Y., Nitta, K., & Kaneko, Y. (2013). Observation of polyphosphate bodies and DNA during the cell division cycle of Synechococcus elongatus PCC 7942. *Plant biology (Stuttgart, Germany)*.

[149] Sharma, N. K., Rai, A. K., & Stal, L. J. (2013). *Cyanobacteria: An Economic Perspective*. John Wiley & Sons.

[150] Shay, J. W. & Wright, W. E. (2000). Hayflick, his limit, and cellular ageing. *Nature reviews. Molecular cell biology*, 1(1), 72–6.

[151] Sherratt, D. J. (2003). Bacterial chromosome dynamics. *Science (New York, N.Y.)*, 301(5634), 780–5.

[152] Shively, J. M., Ball, F., Brown, D. H., & Saunders, R. E. (1973). Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. *Science (New York, N.Y.)*, 182(4112), 584–6.

[153] Skarstad, K., Boye, E., & Steen, H. B. (1986). Timing of initiation of chromosome replication in individual Escherichia coli cells. *The EMBO journal*, 5(7), 1711–7.

[154] Skarstad, K., von Meyenburg, K., Hansen, F. G., & Boye, E. (1988). Coordination of chromosome replication initiation in Escherichia coli: effects of different dnaA alleles. *J. Bacteriol.*, 170(2), 852–858.

[155] Sliusarenko, O., Heinritz, J., Emonet, T., & Jacobs-Wagner, C. (2011). High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Molecular microbiology*, 80(3), 612–27.

[156] Smith, R. M. & Williams, S. B. (2006). Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium Synechococcus elongatus. *Proceedings of the National Academy of Sciences of the United States of America*, 103(22), 8564–9.

[157] Smolensky, M. H. & Peppas, N. A. (2007). Chronobiology, drug delivery, and chronotherapeutics. *Advanced drug delivery reviews*, 59(9-10), 828–51.

[158] Storch, K.-F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F. C., Wong, W. H., & Weitz, C. J. (2002). Extensive and divergent circadian gene expression in liver and heart. *Nature*, 417(6884), 78–83.

[159] Stuurman, N., Amdodaj, N., & Vale, R. (2007). Micro-Manager: Open Source software for light microscope imaging. *Microscopy Today*, 15(3), 42 – 43.

[160] Tabor, J. J., Levskaya, A., & Voigt, C. A. (2011). Multichromatic control of gene expression in Escherichia coli. *Journal of molecular biology*, 405(2), 315–24.

[161] Takai, N., Nakajima, M., Oyama, T., Kito, R., Sugita, C., Sugita, M., Kondo, T., & Iwasaki, H. (2006). A KaiC-associating SasA-RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 103(32), 12109–14.

[162] Tanaka, S., Kerfeld, C. a., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C., & Yeates, T. O. (2008). Atomic-level models of the bacterial carboxysome shell. *Science (New York, N.Y.)*, 319(5866), 1083–6.

[163] Tanaka, S., Sawaya, M. R., Phillips, M., & Yeates, T. O. (2009). Insights from multiple structures of the shell proteins from the beta-carboxysome. *Protein science : a publication of the Protein Society*, 18(1), 108–20.

[164] Thaben, P. F. & Westermark, P. l. O. (2014). Detecting rhythms in time series with RAIN. *Journal of biological rhythms*, 29(6), 391–400.

[165] Thaiss, C. A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A. C., Abramson, L., Katz, M. N., Korem, T., Zmora, N., Kuperman, Y., Biton, I., Gilad, S., Harmelin, A., Shapiro, H., Halpern, Z., Segal, E., & Elinav, E. (2014). Transkingdom Control of Microbiota Diurnal Oscillations Promotes Metabolic Homeostasis. *Cell*, 159(3), 514–29.

[166] Tigges, M., Marquez-Lago, T. T., Stelling, J., & Fussenegger, M. (2009). A tunable synthetic mammalian oscillator. *Nature*, 457(7227), 309–12.

[167] Tomita, J., Nakajima, M., Kondo, T., & Iwasaki, H. (2005). No transcription-translation feedback in circadian rhythm of KaiC phosphorylation. *Science (New York, N.Y.)*, 307(5707), 251–4.

[168] Toro, E. & Shapiro, L. (2010). Bacterial chromosome organization and segregation. *Cold Spring Harbor perspectives in biology*, 2(2), a000349.

[169] Valencia S, J., Bitou, K., Ishii, K., Murakami, R., Morishita, M., Onai, K., Furukawa, Y., Imada, K., Namba, K., & Ishiura, M. (2012). Phase-dependent generation and transmission of time information by the KaiABC circadian clock oscillator through SasA-KaiC interaction in cyanobacteria. *Genes to cells : devoted to molecular & cellular mechanisms*, 17(5), 398–419.

[170] Viollier, P. H., Thanbichler, M., McGrath, P. T., West, L., Meewan, M., McAdams, H. H., & Shapiro, L. (2004). Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, 101(25), 9257–62.

[171] Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., & Jun, S. (2010). Robust growth of Escherichia coli. *Current biology : CB*, 20(12), 1099–103.

[172] Watanabe, S., Ohbayashi, R., Shiwa, Y., Noda, A., Kanesaki, Y., Chibazakura, T., & Yoshikawa, H. (2012). Light-dependent and asynchronous replication of cyanobacterial multi-copy chromosomes. *Molecular Microbiology*, 83(4), 856–865.

[173] Webb, C. D., Teleman, A., Gordon, S., Straight, A., Belmont, A., Lin, D. C.-H., Grossman, A. D., Wright, A., & Losick, R. (1997). Bipolar Localization of the Replication Origin Regions of Chromosomes in Vegetative and Sporulating Cells of B. subtilis. *Cell*, 88(5), 667–674.

[174] Weber, W., Stelling, J., Rimann, M., Keller, B., Daoud-El Baba, M., Weber, C. C., Aubel, D., & Fussenegger, M. (2007). A synthetic time-delay circuit in mammalian cells and mice. *Proceedings of the National Academy of Sciences of the United States of America*, 104(8), 2643–8.

[175] Whitney, S. M., Houtz, R. L., & Alonso, H. (2011). Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco. *Plant physiology*, 155(1), 27–35.

[176] Wijffels, R. H., Kruse, O., & Hellingwerf, K. J. (2013). Potential of industrial biotechnology with cyanobacteria and eukaryotic microalgae. *Current opinion in biotechnology*, 24(3), 405–13.

[177] Wilmotte, A. M. R. & Stam, W. T. (1984). Genetic Relationships among Cyanobacterial Strains Originally Designated as 'Anacystis nidulans' and Some Other Synechococcus Strains. *Microbiology*, 130(10), 2737–2740.

[178] Wirz-Justice, A., Benedetti, F., Berger, M., Lam, R. W., Martiny, K., Terman, M., & Wu, J. C. (2005). Chronotherapeutics (light and wake therapy) in affective disorders. *Psychological medicine*, 35(7), 939–44.

[179] Wohlert, S.-E., Lomovskaya, N., Kulowski, K., Fonstein, L., Occi, J. L., Gewain, K. M., MacNeil, D. J., & Hutchinson, C. (2001). Insights about the biosynthesis of the avermectin deoxysugar L-oleandrose through heterologous expression of Streptomyces avermitilis deoxysugar genes in Streptomyces lividans. *Chemistry & Biology*, 8(7), 681–700.

[180] Wörsdörfer, B., Woycechowsky, K. J., & Hilvert, D. (2011). Directed evolution of a protein container. *Science (New York, N.Y.)*, 331(6017), 589–92.

[181] Wulff, K., Gatti, S., Wettstein, J. G., & Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nature reviews. Neuroscience*, 11(8), 589–99.

[182] Xu, P., Ranganathan, S., Fowler, Z. L., Maranas, C. D., & Koffas, M. A. G. (2011). Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metabolic engineering*, 13(5), 578–87.

[183] Yeates, T. O., Crowley, C. S., & Tanaka, S. (2010). Bacterial microcompartment organelles: protein shell structure and evolution. *Annual review of biophysics*, 39, 185–205.

[184] Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C., & Shively, J. M. (2008). Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature reviews. Microbiology*, 6(9), 681–91.

[185] Zwicker, D., Lubensky, D. K., & ten Wolde, P. R. (2010). Robust circadian clocks from coupled protein-modification and transcription-translation cycles. *Proceedings of the National Academy of Sciences of the United States of America*, 107(52), 22540–5.